



Essays on Education and the Macroeconomy

Vinzenz Johannes Zieseimer

Thesis submitted for assessment with a view to obtaining the degree of
Doctor of Economics of the European University Institute

Florence, 12 November 2018

European University Institute
Department of Economics

Essays on Education and the Macroeconomy

Vinzenz Johannes Zieseimer

Thesis submitted for assessment with a view to obtaining the degree of
Doctor of Economics of the European University Institute

Examining Board

Prof. Árpád Ábrahám, EUI, Supervisor
Prof. Piero Gottardi, EUI and University of Essex
Prof. Bas Jacobs, Erasmus University Rotterdam
Prof. Dirk Krueger, University of Pennsylvania

© Vinzenz Johannes Zieseimer, 2018

No part of this thesis may be copied, reproduced or transmitted without prior
permission of the author



Researcher declaration to accompany the submission of written work

I Vinzenz Johannes Zieseimer certify that I am the author of the work *Essays on Education and the Macroeconomy* I have presented for examination for the PhD thesis at the European University Institute. I also certify that this is solely my own original work, other than where I have clearly indicated, in this declaration and in the thesis, that it is the work of others.

I warrant that I have obtained all the permissions required for using any material from other copyrighted publications.

I certify that this work complies with the *Code of Ethics in Academic Research* issued by the European University Institute (IUE 332/2/10 (CA 297)).

The copyright of this work rests with its author. (Quotation from it is permitted, provided that full acknowledgement is made.) This work may not be reproduced without my prior written consent. This authorisation does not, to the best of my knowledge, infringe the rights of any third party.

Statement of inclusion of previous work:

I confirm that chapter 2 was jointly co-authored with Mr Benedikt Dengler and I contributed 50% of the work.

I confirm that chapter 3 was jointly co-authored with Mr Benedikt Dengler and Mr Árpád Ábrahám and I contributed 33% of the work.

Signature and Date:

Vinzenz Johannes Zieseimer, 6 September 2018

Abstract

This thesis consists of three independent chapters, related by a common theme: the role of education in the macroeconomy.

The first chapter considers the role of higher education policies in intergenerational mobility. Student loans and grants increase the possibilities for low-income students to attend college and earn high incomes later in life. For that reason, they are commonly assumed to increase intergenerational mobility. Instead, the chapter shows that education policies have another effect, working in the opposite direction: they reduce the relative importance of other components of earnings such as luck, while those components are a greater source of mobility. Which of the two effects dominates is an empirical question. To that end, the chapter develops and parameterizes a model of the markets for higher education and labor. The results show a trade-off between welfare and intergenerational mobility.

The second chapter connects two disparate strands of literature on earnings inequality. On the one hand, skill-biased technological change describes how general equilibrium effects between different types of workers shape the income distribution. On the other, the literature on taxation suggests that incentives to accumulate human capital drive the earnings distribution. The chapter combines both approaches, underpinned by an empirical analysis of occupational skill data. It finds that incentive changes in taxation like those that occurred in the second half of the 20th century can lead to polarization of the labor market.

The third chapter really concerns education in *economics*, rather than education in the *economy*. It analyses the completion times of students in top European PhD programs. These are comparable to their counterparts in the United States, with a median that is approaching six years and a higher average. The publication of the present thesis helps counter the trend.

Acknowledgments

I would like to thank my supervisor Árpád Ábrahám. He proved to be an endless and much needed source of enthusiasm and advice. The same thanks are due to my second advisor, Piero Gottardi, who was always there to provide a listening ear.

My work has benefited from the feedback and advice from many faculty and visitors at the EUI. In particular, I would like to mention Philipp Kircher, Dominik Sachs, Axelle Ferriere, and Juan Dolado. I also want to thank Victor Rios-Rull and Dirk Krueger for welcoming me to the University of Pennsylvania as a visiting student.

I wrote the second chapter of this thesis together with Benedikt Dengler, and the third chapter with Benedikt and Árpád. Matic Petricek provided outstanding research assistance on the third chapter.

I gratefully acknowledge financial support by Nuffic, the European University Institute, and the Prins Bernhard Cultuurfonds.

Finally, I would like to thank my wife Tanya, my family, my friends, and my cohort at EUI Economics. Thank you for your patience, tolerance, and company.

Contents

1	Higher Education Policies and Intergenerational Mobility	9
1.1	Introduction	9
1.1.1	Overview	9
1.1.2	Literature	15
1.2	Decomposing Intergenerational Mobility	16
1.3	A Theory of Intergenerational Earnings Persistence	18
1.3.1	A Model of Colleges	19
1.3.2	A Model of the Labor Market	21
1.4	Parameterization	31
1.4.1	Estimation	32
1.4.2	Moment Matching	37
1.5	Implications	43
1.5.1	Intergenerational Mobility	43
1.5.2	Heterogeneity of Intergenerational Mobility	44
1.5.3	College Entry and Heterogeneity	46
1.5.4	Paying for College	49
1.5.5	Labor Earnings	51
1.6	Counterfactuals	53
1.6.1	A Decomposition of Earnings Persistence	53

1.6.2	Removing Borrowing Constraints	55
1.6.3	Exploring the Effect of Education Policies	57
1.7	Conclusion	59
1.8	Appendix	61
1.8.1	Computation	61
1.8.2	Student Aid in the United States	62
2	Polarization: A Supply-Side Mechanism	65
2.1	Introduction	65
2.1.1	Motivation	65
2.2	Manual and Cognitive Skills	71
2.2.1	Data	71
2.2.2	Empirics	74
2.3	Model	79
2.3.1	Environment	79
2.3.2	Individual Problem	80
2.4	Tax Policy, Inequality Changes and Polarization	81
2.4.1	A simple static framework	81
2.4.2	Policy changes and earnings distribution	82
2.4.3	Conditions for Inequality Growth and Polarization	84
2.5	Models versus Data	87
2.5.1	Across Countries	87
2.5.2	Over Time	89
2.6	An Enriched Model	90
2.6.1	Model Description	90
2.6.2	Parameterization	92

2.6.3	Results	94
2.6.4	Sensitivity Analysis	96
2.7	Conclusion	98
2.8	Appendix	100
2.8.1	PCA Results	100
3	Economics PhD Programs in Europe: Completion Times and More	103
3.1	Introduction	103
3.2	Economics PhD Programs in Europe	104
3.3	Data Collection and Processing	106
3.4	Sample Selection and Representativeness	107
3.5	Results	108
3.6	Conclusion	110
3.7	Appendix	112
3.7.1	Tables and Figures	112
	Bibliography	132

Chapter 1

Higher Education Policies and Intergenerational Mobility

1.1 Introduction

1.1.1 Overview

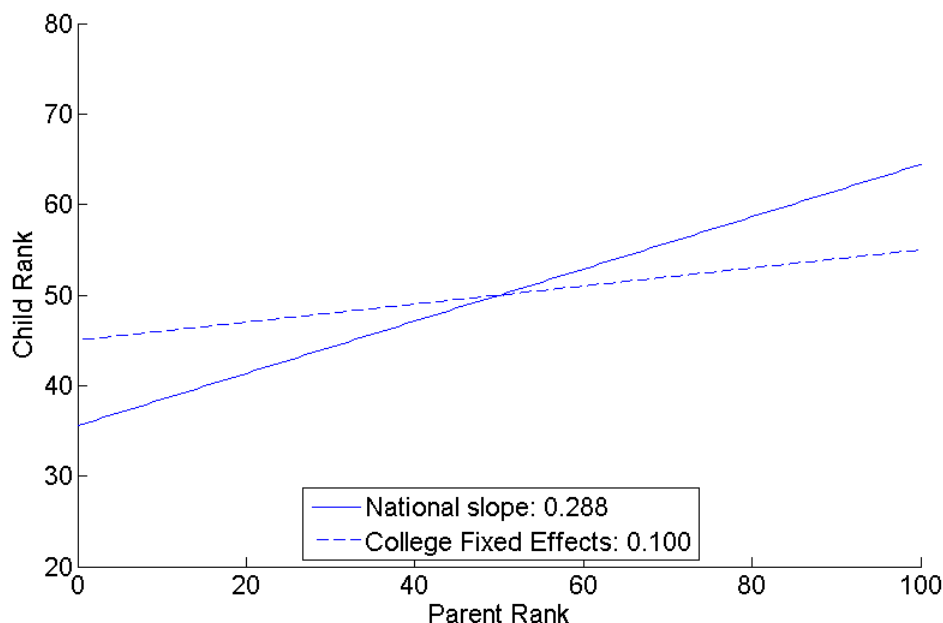
Intergenerational mobility is considered a key macroeconomic variable. It measures how much children differ from their parents, on average. This tells us to what extent one's outcomes in life depend on circumstances at birth. Intergenerational mobility in economic terms is typically measured by labor earnings. The less children's earnings are related to those of their parents, the more mobile a society is.

Recent empirical work suggests that higher education is key to understanding the causes of intergenerational mobility. Chetty, Friedman, Saez, Turner, and Yagan (2017) find that there is a strong correlation between parental earnings and child earnings for the United States as a whole (a rank-rank correlation of 0.288). But this correlation is much smaller when including college fixed effects (0.100), i.e. a child's earnings are almost unrelated to that of their parent once we know what college the child goes to. This is found to be true irrespective of the type of college. Figure 1.1 illustrates the analysis.¹ The findings by Chetty et al. add to

¹These are measures of persistence of earnings across generations, which are the inverse of mobility: the lower intergenerational persistence, the higher intergenerational mobility. Chetty et al. combine data from federal income tax returns and from the Department of Education to link information on the earnings of two generations to their educational choices, and the characteristics of colleges they attend. These data are available for individuals from the 1980-1982 birth cohorts. In 2014, the time of the last earnings measurement, children of those cohorts were in their early 30s, at which point measures of intergenerational earnings persistence typically stabilize. Parental income is defined as average parental earnings when the children are

an established literature that suggests human capital is the main culprit in intergenerational mobility, and that once we understand educational outcomes we will largely understand intergenerational earnings persistence.

Figure 1.1: Controlling for Colleges



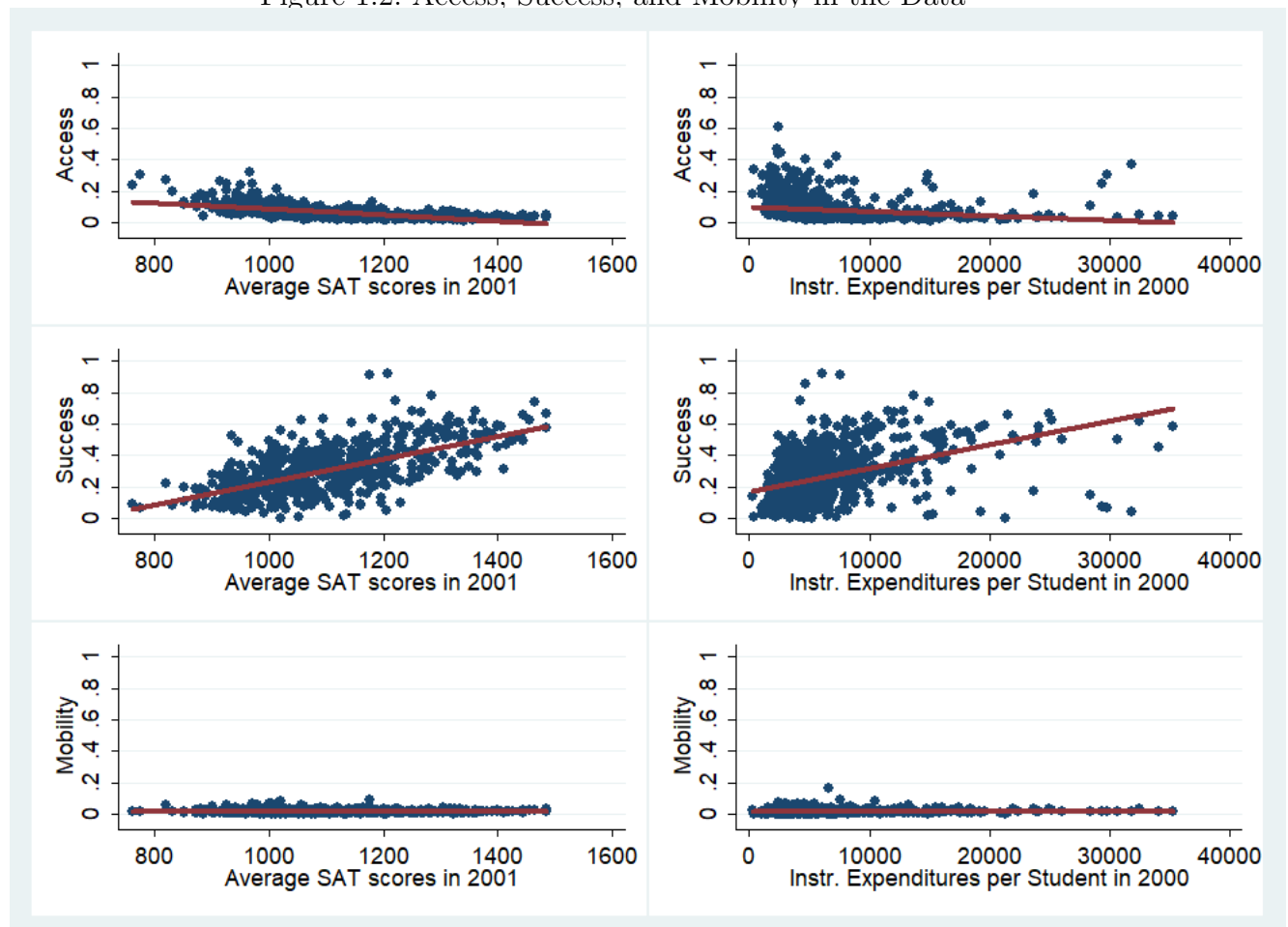
Graphs are created by fitting a straight line with slope equal to the IGE estimate through the point (50,50).

Chetty et al. also document intergenerational mobility by type of college. Specifically, they calculate the proportion of students in each college that comes from a low income background, defined as a family income in the bottom 20%. This is then seen as a measure of *access* to the college. Next, they find the proportion of students from a low income family that reach the top 20% of the income distribution, which they take as a measure of a college's *success* in generating mobility. Finally, the product of these two is the share of students in a college that comes from a bottom 20% family and reaches the top 20%. This is a measure of *mobility*. Figure 1.2 displays these measures per college, first by a measure of average student ability on the left, and then by a measure of educational investment on the right. To emphasize the absolute value of the probability measures, all graphs have a common vertical axis. Access decreases as a college enrolls students with higher average SAT scores, or as it spends more on instructional expenditures. At the same time, the success of those students that do attend

aged 15 to 19, and child income is measured over the year 2014. The college a student was enrolled in longest counts as the college that the student attended. Their preferred measure of intergenerational persistence are rank-rank regressions of parental and child earnings. (Both parents and children are assigned a rank within their own distribution. Then, child rank is regressed on parental rank. This procedure is the same as the Spearman correlation, but additionally allows for the inclusion of controls, such as college attended.)

increases. As a result, mobility across colleges is remarkably flat and low. In short, college heterogeneity appears to be an important part of the story. These findings also suggest that higher education policies may have an important role to play: policies help low-income students pay for college, and college expenditures seem to propel them upward in the earnings distribution.

Figure 1.2: Access, Success, and Mobility in the Data



Each point represents a college. Linear fit included for emphasis. Only 4-year private colleges are included, and those who report instructional expenditures per student above \$40,000 are excluded as outliers.

Data from Chetty et al. (2017).

Building on these insights, this paper develops a theory of the persistence of labor earnings across generations, paying special attention for the role of higher education. The theory is parameterized to represent the United States economy in the early 2000s, and used to explore the relationship between higher education policies and intergenerational mobility.

The main result is the following: higher education policies, such as student grants and loans, actually increase the persistence of earnings, thereby reducing intergenerational mobility.

To understand why this may be so, we start by looking at the sources of labor earnings. The macroeconomic literature on labor earnings identifies two main sources of gross wages. First, wages depend on one's skills, often referred to as *human capital*. Human capital is part innate ability, and part learned through education or on the job. Second, wages also depend on idiosyncratic shocks to income: careers do not always follow straight lines, and neither do the earnings derived from them.

The common assumption that policies increase mobility revolves around human capital. Building human capital costs money that the children of the poor may not have. Education policies pay for expensive colleges, therefore allowing smart children from poor families to earn more than their parents.

But this story ignores the role of luck. Human capital tends to persist over generations, as children inherit their parent's ability. If earnings were only due to human capital, earnings would be much more persistent than what we observe in data. Idiosyncratic shocks to income, on the other hand, mix up outcomes. The relative importance of these two components, together with their level, determines overall mobility.

So why do policies reduce the intergenerational mobility of earnings? By enabling human capital investment, they make human capital a relatively more important component of earnings, and reduce the role of randomness. Put simply, they turn an economy driven by luck into an economy driven by ability. The latter economy, however, exhibits much less intergenerational mobility.

Key to this result is to establish the relative importance of the two competing channels, and the effect higher education policies have on them. That requires a model that can generate counter-factual states of the world. This paper builds such a model. The below outlines how I ensure that all of the model's effect sizes are driven by relevant data or previous empirical research.

First, the model needs to take into account the behavior of institutions of higher education. To this end, the paper describes a theory of competitive colleges that translate spending in human capital investment one-to-one, and argues why this view of colleges is appropriate for the question at hand.

Second, it needs to account for the effectiveness of higher education. At the core of this lie questions of causality: how important is time and money invested in higher education versus ability? How does this differ by college selectiveness? And how much value does college add versus going to work straight away? Quasi-experimental results on college effectiveness from Hoxby (2016a) help to discipline the model in this respect.

Third, it needs to make sense of who joins which institution, which means taking into account, amongst other matters, financial constraints and government policies. Therefore, public colleges, grants, and student loans are all modeled. The availability of grants is directly estimated from survey data. Regarding student loans, the paper relies on previous work by a number of authors who study financial constraints in college financing. In particular, the paper follows Abbott et al. (2013) in the sources of financing it considers and its model of the US student loan system.

Fourth, because earnings persistence is measured at a later point in life, it needs to account for different components of earnings and their path over the adult life cycle. Here, the paper follows the approach taken by Huggett, Ventura, and Yaron (2011). It uses a human capital production function to match the hump-shaped life-cycle of earnings, as well as the growing variance of earnings across members of a cohort as they grow older. It also adds idiosyncratic income shocks and a labor-leisure choice to human capital. Extensive literatures exist on either of these two components of earnings, which are used to quantify their dynamics and importance.

Fifth and final, links between generations need to be accounted for. These links come in three parts. First, test scores of matched pairs of parents and children are used to quantify the intergenerational persistence of learning ability. The paper does not advocate that approach *per se*, although that has been done elsewhere. Instead it rests at demonstrating that the overall model makes accurate predictions with regards to the intergenerational persistence of earnings. Notably, this is so without targeting these figures at any point in the parameterization of the model. Second, inter-vivos transfers from parents to children which help to finance college can be observed in the data. They mostly appear to take place on the basis of need, and are therefore modeled and quantified as such. Third, grants and loans are sometimes dependent on family income status. Such grants and loans are modeled explicitly.

A credible parameterization of the model is only possible due to recent empirical advances. Two pieces of evidence stand out in particular, both combining large sets of data from administrative sources. Work by Chetty et al. (2017) provides several results on mobility by colleges which serve to inform the role of higher education more generally. Hoxby (2016a) uses discontinuities at admissions thresholds to infer the causal effect of higher education spending on earnings later in life. Because she uses administrative data on applications, admissions, and earnings, her work provides these results for the entire spectrum of colleges. These findings are key ingredients to the present work, which combines them with further data on higher education and the labor market.

It has already been mentioned that the the parameterization of the model does not target any measures of intergenerational persistence, so that these can be used to assess the model's

validity. We will also see that the model presented in this paper reproduces the patterns found by Chetty et al. Further to that, the model also has a number of implications regarding college entry, college heterogeneity, the sources of college financing, and several components of labor earnings, none of which are targeted as part of the parameterization of the model. The paper compares these implications to data, and argues that the model represents these aspects well. In the process, the paper shows that a human capital-based model can explain the facts on intergenerational mobility.

The main result of this paper has several implications. To begin, it shows a direct trade-off between classical welfare measures and intergenerational mobility. As a result, it may not be wise for policy makers to use higher education policies to target intergenerational mobility. Next, one should be careful when comparing countries, regions, or periods by measures of intergenerational mobility. Competing channels are at work, so that one cannot infer the quality of policies from these measures. Finally, it may be more informative to analyze the effects of policies on the components of earnings separately. For example by measuring mobility in educational attainment rather than in earnings. The last two points have already been taken up in related empirical research. Landersø and Heckman (2017) show that while intergenerational mobility is larger in Denmark than in the United States when measured by earnings, the same is not true when measuring mobility by educational attainment.

This paper includes two further sets of results. The first is on the role of college heterogeneity and financial constraints. It turns out that even when students are essentially unconstrained in their extensive margin of college choice (going or not), they may still be constrained in their intensive margin of college choice (which college to go to). Thus, accounting for college heterogeneity is crucial in understanding the welfare implications of higher education policies.

Second, the model yields a decomposition of the current persistence of earnings across generations. Roughly half of earnings persistence is determined before the start of adult life. Thus, while it is true that childhood is perhaps most important in determining intergenerational mobility, higher education and adult life are worth studying. Of the remaining persistence, about a third is due to money from parents, and two thirds to government policies. This confirms previous results from similar quantitative theoretical work on the determinants of college enrollment: parental resources are an important source of college financing, and may be very responsive when education policies change. They thereby greatly reduce the impact of education policies. Nevertheless, education policies have significant impact.

The remainder of this section discusses related literature (1.1.2). Section 1.2 uses theory to demonstrate why higher education policies have an ambiguous effect on intergenerational mobility. Section 1.3 contains a full description of the paper's theory. Section 1.4 describes

how the parameters of the resulting economic model are either estimated or set to match moments. Section 1.5 explores the ability of the model to match aspects of the data that were not targeted in the parameterization of the model. It also discusses what we can learn from the theory's positive implications. Section 1.6 contains results from counterfactual policies. Section 2.7 concludes.

1.1.2 Literature

There is a macroeconomic literature that connects education policies to intergenerational mobility. Lee and Seshadri (2014) argue that a rich life-cycle model with intergenerational links explains a number of intergenerational relationships well, in particular the intergenerational elasticity of earnings. They focus more on development of human capital during childhood, and less on college heterogeneity. In particular, they do not allow for heterogeneous spending on higher education.

Holter (2015) similarly builds a quantitative model of intergenerational mobility. He then investigates the extent to which differences in tax and education policies can explain cross-country differences in intergenerational mobility. Herrington (2015) also looks at the effect of taxes and education policies on inequality and intergenerational mobility through the lens of a model, comparing policies of the United States to those of Norway. Kotera and Seshadri (2017) analyze the effect of public school spending at the compulsory stage on regional variation in intergenerational mobility. The current paper features a richer model of inter-generational links, and takes a more granular look at higher education policies. It is also the first to discuss the theoretically ambiguous effect of these policies.

Positive implications of education policies on college enrollment are studied by a number of authors. Important work is by Lochner and Monge-Naranjo (2011b), who consider the effect of student loan policies on the college entry decision of youth that is heterogeneous in ability and family income. Abbott et al. (2013) study the decision to go to college or not in a quantitative model with intergenerational transfers, and find that these transfers are an important adjustment margin that dampen the effects of education policies in equilibrium. They only consider one type of college. Empirical work on the incidence of financial constrainedness is summarized in Lochner and Monge-Naranjo (2011a), who find increased evidence for such incidence in recent years. None of these papers focus on intergenerational mobility.

There is a large normative literature on education policies, often in combination with taxation. While these papers answer a different question, their insights guide the discussions of policy optimality in this work. Krueger and Ludwig (2016) focus is on optimal taxation with (almost) linear instruments. Their paper also includes intergenerational transfers, but only has one type of college, and considers general equilibrium effects as well as the importance of

the transition between different policy regimes. Bovenberg and Jacobs (2005, 2011) find that while education subsidies themselves distribute resources to the well to do, their optimal level may still be positively related to tax rates. This is because they undo the disincentive effects of taxation on human capital formation. The same issue has been studied in a dynamic theoretical framework by Stantcheva (2017), and in a quantitative framework by Hanushek, Leung, and Yilmaz (2003). Further, in an incomplete market where students cannot borrow against future income, there is a role for government-provided student loans. These are studied in a dynamic framework by Findeisen and Sachs (2016a). Finally, Findeisen and Sachs (2016b) study an economy with the same motivation for education subsidies as in Bovenberg and Jacobs, but with an extensive margin for college choice and under financial constraints. In that case, the government wants to efficiently target those who would optimally be students from a social standpoint, but who would not enter college in the absence of policy intervention. It can do so by need-dependent grants, essentially using parental income as a tag of financial constraints. Insights from this normative literature guide some of the discussion of policy in this paper.

Holmlund, Lindahl, and Plug (2011) attempt to synthesize a growing literature on the effect of parents' schooling on children's schooling. Overall, the causal effect of changes to parents' schooling on children's schooling appears to be small relative to the total correlation between parents' schooling and that of their children. That is in line with this paper's assumptions, since it takes ability at the start of adult life (and its transition across generations) as given. At the same time, Holmlund et al. find that the causal effect of parental schooling may be a sizable part of the 'nurture' component to the overall correlation. The mechanism by which this occurs is unclear: it may be due to parents' schooling leading to higher incomes, or due to other factors. The former is modeled in this paper where it concerns tertiary education. Incorporating the effects of policies on the development of children in the earlier stages of the life cycle would likely increase the effects sizes reported in this paper. This is discussed further in subsection 1.4.1.

1.2 Decomposing Intergenerational Mobility

The most commonly used measure of intergenerational persistence is the intergenerational elasticity of earnings (IGE), measured as β^{IGE} in the regression equation below:

$$\log(y') = \beta_0 + \beta^{IGE} \log(y) + \epsilon \quad (1.1)$$

Here, y is a measure of parental earnings, and y' measures the earnings of their children. As we will see later, measurements of β^{IGE} in the literature have a wide range between 0.3 and 0.6, suggesting that a 1% increase in parental earnings is expected to lead to 0.3% to 0.6%

higher earnings for children. In other words, earnings are persistent over generations but not perfectly: they regress to the mean.

As is well known, common estimators of the equation above (such as OLS) are unbiased estimators of

$$\beta^{IGE} = \frac{Cov(\log y', \log y)}{Var(\log y)}.$$

When an economy is in steady state, $Var(\log y) = \sqrt{Var(\log y)}\sqrt{Var(\log y')}$, so that the IGE measures $Cor(\log y', \log y)$.

In a typical macroeconomic model of labor, wages would be represented by human capital (h) times some idiosyncratic shock to income (x). The latter represents any form of luck not related to ability or education. Using this as our measure of earnings (and abstracting from labor supply for this exposition), we have:

$$\log y = \log h + \log x.$$

If we now (in addition to considering steady state economies) assume that luck is entirely independent from human capital as well as from parental characteristics, we can write:

$$\begin{aligned} \beta^{IGE} &= \frac{Cov(\log h, \log h') + Cov(\log x, \log h')}{Var(\log h) + Var(\log x)} \\ &= \frac{Cor(\log h, \log h')Var(\log h) + \left[Cor(\log x, \log h') \frac{\sqrt{Var(\log h)}}{\sqrt{Var(\log x)}}\right] Var(\log x)}{Var(\log h) + Var(\log x)} \end{aligned} \quad (1.2)$$

Now, the IGE is a weighed mean of the correlation between two generations' human capital on the one hand, and a measure of the influence of parental luck on children's human capital on the other. The respective weights are the variance of log human capital, and the variance of log income shocks.

The economic role of higher education policies is to relieve financial constraints. So how does the IGE change when financial constraints are relieved? Table 1.1 makes a comparison using the components of the expression above.

The first component is conventional: the correlation of human capital across generations. Financial constraints in education deny the children of poor parents the education they need to go to college, thereby keeping them poor. This increases the persistence of earnings across generations, as poor parents are now more likely to produce poor children (compared to the unconstrained case). This is indeed true, but applies to the intergenerational correlation of human capital only (row 1 of Table 1.1).

The second component of our weighted sum goes the same way. Parental luck increases children's human capital when they are financially constrained, but less so in the unconstrained

case: without constraints, children’s potential outcomes do not depend on the financial situation of their parents. Thus, the second term (row 2 of Table 1.1) also increases due to financial constraints.

So how can the role of education policies be ambiguous? It turns out that the ‘weights’ given to the two terms provide a sharp trade-off (rows 3 and 4 of Table 1.1). As I will argue later, the variance of log income shock is unlikely to vary much due to education policies. However, the variance of log human capital may. If the same education policies that release constraints also increase the variance of human capital in the system, then that makes the correlation in human capital more important (and luck less so). Because we would expect the human capital of two generations to be more correlated than parental luck and children’s human capital (the first term dominates the second), it is not clear which way the IGE will finally move - that becomes a matter of measurement.

Table 1.1: The effect of financial constraints on the IGE

	Unconstrained	Constrained
$Cor(\log h, \log h')$		$<$
$\left[Cor(\log x, \log h') \frac{\sqrt{Var(\log h)}}{\sqrt{Var(\log x)}} \right]$		$<$
$Var(\log x)$		$=$
$Var(\log h)$		$>$

In fact, in the remainder of this paper I find that the variance effect (row 4 of Table 1.1) dominates for the model’s equivalent of current US higher education policies. As a result, education policies actually decrease intergenerational mobility. The paper also demonstrates that the result does not depend on the measure of mobility I discuss here.

1.3 A Theory of Intergenerational Earnings Persistence

This paper’s theory of intergenerational persistence focuses on higher education. Below, the first subsection sketches a stylized model of colleges in a competitive setting.² As a result of that model, students can choose how much to invest in their own education, with colleges just translating spending into investment.

In the second subsection, the same model of colleges and college choice is embedded in a model of the labor market. Higher education takes place in the first period at the start of adulthood. Thereafter, agents go through a life cycle of earnings and related choices, and

²The terms ‘college’ and ‘institution of higher education’ are used interchangeably. Later, the model will be brought to the data in such manner that the higher education phase represents the entire higher education career.

have children of their own. Together, these models then describe earnings persistence and the role of higher education in it. Modeling choices are highlighted as they appear. The individual's decision problem is specified in full. The subsection ends with a definition of stationary equilibrium and a description of the solution method.

1.3.1 A Model of Colleges

Students are defined by their learning ability α . When going to a college, that learning ability combines with time spent studying e and money invested in education d to form human capital $h(\alpha, d, e)$. To have money invested in education, the student must go to a college, which charges price $\tilde{d}(d, q, \alpha)$ for an investment of d .³ In principle, the college can condition that price on the student's parental income status q (so that the price of college becomes need based) and on the ability α of the student (which makes the price merit based).

The student's decision making problem is discussed in detail in the following subsection. For now, it suffices to say that the student chooses from available colleges based on the price he must pay for d , since that is the only thing the college has to offer. Peer effects, whether through learn or networking, as well as the signaling value of going to a college are not modeled explicitly. The model will be parameterized to match the actual earnings returns to investing in education, so that it does not matter for the purposes of this paper whether these returns are due to actual learning or other sources.

Private colleges are indexed by their level of educational spending per student d , which is the same for each student in the college. Colleges do not face any fixed cost. Instead, they have access to an education technology in which they simply incur the cost of educational investment for each student.⁴ New colleges can freely enter any market for a d type college, free of cost. Suppose they either value profits (for-profit colleges) or their existence (not-for-profit colleges). Then we have the following:

$$\forall q \forall \alpha \quad \tilde{d}(d, q, \alpha) \leq d.$$

Any type of student receives an educational investment that is at least as large as their spending on college. If this condition were violated for any type of student, new colleges would enter and offer the same services at a lower price until some college offers $\tilde{d}(d, q, \alpha) = d$.

³Price here is meant to refer to the price for college, and not for food, lodging, and the like. Those are considered consumption items for the purposes of this paper, and will be treated as such when connecting theory and data.

⁴This is a simplification that does not come at much of a cost. Because costs now scale linearly in the number of students, the number of students in each college will be indeterminate. This is not important for the purposes of this paper.

Can colleges exist for whom $\tilde{d}(d, q, \alpha) < d$ for some type of student? Yes, if they have other income that they choose to invest in their students' education. A clear example of such income would be endowment income. In any case, due to the result above, any college pricing schedules can always be written as follows:

$$\tilde{d}(d, q, \alpha) = d - g^I(q, \alpha), \quad \text{where } g^I(q, \alpha) \geq 0.$$

This type of pricing schedule is precisely what we observe in data on college pricing. Colleges typically post a sticker price, from which they offer discounts in the form of explicit institutional student grants. Because of this practice, we can separately observe the discounts in data on institutional student grants. As a result, we do not need to make assumptions on colleges' objective functions for the purposes of this paper.

The result of all of the above is a model of 'translated spending'. Students decide what to spend on higher education, and through competitive colleges that same amount (plus any grants received) is invested in their human capital. This way of thinking about goods investment in human capital is much in line with the macroeconomic literature. Some papers in other literatures explicitly model the behavior of colleges. For example, Epple, Romano, and Sieg (2006) analyze a model with quality maximizing colleges, peer-effects and a preference for low-income student enrollment, where price discrimination leads to student sorting over colleges. Epple, Romano, Sarpça, and Sieg (2013) then adapt this framework to include public universities and endow students with idiosyncratic preferences over colleges. For a number of reasons, these frameworks are less applicable to questions this papers asks. All these papers take colleges as given, both in numbers and in terms of characteristics, and can therefore not explain why we see the colleges that we see. The strategic interactions these models describe typically become less relevant when the number of colleges grows large - and the number of colleges is large in reality. Idiosyncratic preferences by students over colleges can maintain college pricing power even then, but it is not clear whether this is an empirically relevant channel. Lastly, this work's model makes a sharp prediction on the shape of the pricing function, which other papers have to impose by assumption instead.⁵

Not being explicit about colleges' objectives comes at a cost. How do colleges respond when government education policies adjust? This work will maintain the assumption that they simply do not. Objectives from which this would result are thinkable, although they would be non-standard.

At last, students can in practice choose to enter public colleges. Public colleges in the United States largely function like private ones, with the qualification that the government sets

⁵Recent work by Cai and Heathcote (2018) is an important exception. Cai and Heathcote also model a competitive market for colleges, resulting in an endogenous distribution of colleges. Going beyond this paper, they treat colleges as a 'club good' so that there is a strategic aspect to college choice in their model.

pricing schedules and determines how much money a public college has to spend on education. Availability of public colleges often depends on place of residence. To capture all this, I model one representative public college with its own pricing schedule (also consisting of a sticker price and institutional grants) that is set by the government, to which all students have access. In practice, there is heterogeneity in public colleges, although offerings are set by local governments. This likely makes availability less responsive to demand than private college offerings (as well as dependent on geography). Modeling a single representative public college also greatly reduces computational complexity.

1.3.2 A Model of the Labor Market

This paper considers stationary equilibria. There is a continuum of agents with mass one. Each agent spawns a new agent with a mass identical to its own. We refer to the former as parents and the later as children. The timing of the life-cycle is deterministic and equal for all households. The symbol $'$ is used to denote variables pertaining to an agent's children.

Each agent goes through a life-cycle from age 0 to age T , representing his working life. There are two special phases in this life cycle: In the first period of his life, the agent has access to colleges and, if he chooses to enroll in a college, a system of student grants and loans. At a later point in life (age t^I) he makes an inter-vivos transfer to his children, who begin their life-cycle in the following period. The agent does so because he values the child's expected discounted lifetime utility (at a rate potentially lesser than its own, so that these parents are said to be imperfectly altruistic). This is in line with the literature on inter-vivos transfers, which finds that transfers depend on need or effectiveness (Gale and Scholz, 1994). Modeling the entire life-cycle is useful for at least two reasons: First, several measurements in the empirical literature are taken at specific ages, so that having a model counterpart to these ages is important. Second, the model is then able to capture the life cycle of earnings as in the data, thereby ensuring that returns to education are adequately captured.

In each period, an agent can use his time to work, enjoy leisure, or to invest in human capital. In any period, he can use his resources to consume, save, and repay student loans. At age 0, he chooses whether to go to college or not, and if so how much to invest in a college education. At age t^I he can make an inter-vivos transfer. All resources are expressed in terms of consumption, which is also the model's numeraire. Markets in the model are incomplete in the sense of a Bewley-Aiyagari-Huggett model: agents face idiosyncratic income risk that they cannot insure against. They can borrow using (student) loans and save using a risk-free asset, but face borrowing constraints that potentially constrain their consumption and human capital investment. Individual gross earnings are a combination of human capital and its price, hours worked, and the realization of idiosyncratic wage uncertainty.

Compared to most models of college choice, human capital is continuous in this paper. This allows the model to capture the full effect of education and policies, rather than just the effect on those at the margin of college entry. In college, human capital growth is formed by a constant elasticity production function in ability, goods and time investment. I later show that this functional form captures returns to college well. Those who choose not to go to college or are no longer in college accumulate human capital by a function that is of the Ben-Porath (1967) type, taking only time as an input. This functional form has proven to be successful at capturing the life-cycle of earnings, as well as its heterogeneity across the earnings distribution.

Throughout, agents are assumed to be fully aware of their own ability, which is in line with the finding in the literature that students' uncertainty about their own learning ability is small (cf. Hendricks and Leukhina, 2017). Only one period of fixed length (which will later be set to four years in the data) is used to represent the entire higher education career. Human capital accumulates at the end of that period. This is somewhat restrictive with regards to the time taken to complete college. In reality, some students go to two year colleges, some engage in graduate studies, and so forth. However, it deserves emphasis to say that these different sizes of educational investment are not ruled out: during the period, students can still spend different levels of money and time. The issue is treated with care when connecting the model to data. A similar point holds with regards to drop-outs: these are not modeled explicitly, but that does not undo the empirical strategy of this paper. All relevant data used are conditioned on college entry only.

Each agent in the model economy is linked to their parents in three ways. First, agent's ability to accumulate human capital is correlated with that of their parents. Second, parents endogenously decide how much financial resources to make available to their children as they make initial decisions on human capital investment. Third, government education policies are dependent on parental income. These mechanisms are important in assessing the impact of education policies on human capital investment decisions: when policies change and make more or less resources available, parental transfers are a major compensating margin. And the more persistent ability is across generations, the more correlated wealth and ability will be, reducing the influence of education policies.

The economy contains detailed features of the policy environment in the United States, in particular: taxes, educational subsidies and grants, and student loans: average labor tax rates are non-linear and based on the US tax code, as are other taxes. Section 1.8.2 provides a detailed overview of student aid in the United States in 2003, the year to which the model will be calibrated. The Stafford loan system is explicitly modeled in this paper. To capture subsidies and grants from institutions and all levels of government, the model employs a

flexible specification that allows estimation of these items directly from the data. Finally, students can also choose to go to a representative public college.

Individual's problem

Let s_t denote the stochastic state of the agent's life-cycle at age t , and s^t a history of stochastic states up to age t : $s^t = [s_t, s_{t-1}, \dots, s_1, s_0]$. These histories are suppressed in most of the below, but made explicit where the arguments of the maximization problem are listed.

In the below, c denotes consumption, l leisure, e time investment in human capital, d goods investment in human capital, a assets, b student loans, and v inter-vivos transfers. k denotes college choice (*work*: $k = 1$; *study at a private college*: $k = 2$; *study at a public college*: $k = 3$). For a generic variable x , $I_{[x]}$ is an indicator function that equals one when x is true and zero otherwise. q denotes gross parental wages, which is described in further detail below. The same goes for student loan repayment functions $\pi(b)$ and borrowing constraints. \mathbb{E} is the usual expectations operator. Denote a vector of control variables as follows:

$$\mathbf{z}_t = [c_t, l_t, e_t, a_{t+1}].$$

The initial problem now consists of a college choice, meaning an individual can choose to go to college or not. If the individual does go to a private college, there is an additional choice of the level of educational investment d (which is available at any positive level). If the individual goes to a public college, educational investment d^g is set by the government (as is its price \tilde{d}^g). The individual's choice will depend on a fixed learning ability α , parental wages q (to be discussed below), and their initial asset holdings a_0 . Formally:

$$V(\alpha, q, a_0) = \max_{\{\text{work, study}\}} \left\{ W_0(\alpha, h_0(\alpha_0), 0, a_0), \max_{\{\text{public, private}\}} \left\{ C^g(\alpha, q, a_0), \max_{d \geq 0} C_d(\alpha, q, a_0) \right\} \right\}$$

College enrollment lasts for one period of the model, during which the problem of an indi-

vidual who goes to college d looks as follows:

$$C_d(\alpha, q, a_0) = \max_{\mathbf{z}_0(\bar{s}), b_1} \left\{ \frac{(c_0^\nu l_0^{1-\nu})^{1-\sigma}}{1-\sigma} + W_1(\alpha, h_1(d, e_0, \alpha), \pi_1(b_1), b_1) \right\}$$

subject to:

$$\begin{aligned} c_0(1 + \tau_c) &\leq (1 - l_0 - e_0)wh_0x(\bar{s})(1 - \tau_n(\cdot)) - \tilde{d}(d, q, \alpha) \\ &\quad + a_0(1 + r(1 - \tau_a)) - a_1 - b_1 \end{aligned}$$

$$\begin{aligned} c_0 &\geq 0, \quad 0 \leq l_0, e_0 \leq 1, \quad l_0 + e_0 \leq 1, \\ a_1 &\geq 0, \quad 0 \geq b_1 \geq -\underline{b}_0, \quad a_1 b_1 \geq 0, \quad s_0 = \bar{s}. \end{aligned}$$

Leisure and consumption enter periodic utility multiplicatively. The utility function is tied down by parameters σ and ν .⁶ Consumption and consumption taxes are paid for by what remains of net labor earnings, assets, and student loans after paying for college. Labor earnings are composed of hours worked $(1 - l - e)$, wage rate w , human capital h , and idiosyncratic shock $x(s)$. The idiosyncratic shock $x(s)$ is a function of stochastic state s .

The problem for those who enter the representative public college is the same, only that their education now costs $\tilde{d}^g(d, q, \alpha)$ and yields an investment of d^g .

An individual who does not go to college or has finished studying enters the labor market. The problem of working life is the following:

$$\begin{aligned} W_j(\alpha, h_j, \pi_j, a_j) = & \max_{\left\{ \left\{ \mathbf{z}(s^t) \right\}_{s^t} \right\}_{t=j}^{t^I-1}, \left\{ \left\{ \mathbf{z}(s^t) \right\}_{s^t} \right\}_{t=t^I+1}^{T-1}, \left\{ \left\{ \mathbf{z}(s^{t^I}, \alpha') \right\}_{s^t} \right\}_{\alpha'}, \left\{ \left\{ v(s^{t^I}, \alpha') \right\}_{s^t} \right\}_{\alpha'}} \mathbb{E} \left\{ \sum_{t=j}^{T-1} \beta^t \frac{(c_t^\nu l_t^{1-\nu})^{1-\sigma}}{1-\sigma} + \omega \beta^{t^I} V(\alpha', q', v) \right\} \end{aligned}$$

subject to $\forall t \in \{j, \dots, T-1\}$:

$$\begin{aligned} c_t(1 + \tau_c) &\leq (1 - l_t - e_t)wh_t(d_{t-1}, e_{t-1}, h_{t-1}, \alpha)x(s_t)(1 - \tau_n(\cdot)) \\ &\quad - vI_{[t=t^I]} + a_t(1 + r(1 - \tau_a)) - a_{t+1} - \pi_j \end{aligned}$$

$$\begin{aligned} c_t, v &\geq 0, \quad 0 \leq l_t, e_t \leq 1, \quad l_t + e_t \leq 1, \quad a_{T-1} \geq 0, \quad a_1 \geq 0, \\ q' &= wh_{t^I}x_{t^I}, \quad \alpha' \sim \Gamma_\alpha(\alpha, \alpha'), \quad s_0 = \bar{s}, \quad s_{t+1} \sim \Gamma_s(s_t, s_{t+1}). \end{aligned}$$

⁶With this functional form, the elasticity of inter-temporal substitution is given by $\frac{1}{1-\nu(1-\sigma)}$, and the Frisch elasticity by $\frac{1-\nu(1-\sigma)}{\sigma} \frac{l}{(1-l)}$.

This is a typical life-cycle problem, where next-period utility is discounted by β . The parameter ω discounts the value function of the child's adult life at t^I , which starts at t^{I+1} . At t^I , parents can make an inter-vivos transfer v that affects their child's initial value function. Consumption is paid for using net labor earnings and assets after student loan repayment π .

Human capital So what does an individual gain from college or time spent learning? Both increase human capital, but in different ways. Out of college, human capital production similarly follows from a Ben-Porath (1967) function. This functional form, which has been of much use in the macroeconomics literature, can match the life cycle of earnings well given the right parameterization. Key is that the time input is measured in *human capital hours*, which ensures that hours spent learning or earning are always in direct trade-off:

$$h_{t+1} = h_t(1 - \delta_h) + \alpha(e_t h_t)^{\beta^w}. \quad (1.3)$$

In college, the post-depreciation gain in human capital (denoted $\Delta_\delta h_1 \equiv h_1 - h_0(1 - \delta_h)$) is assumed to have a constant elasticity in both goods and hours of human capital invested, as well as in ability. Combined with the assumption that a zero investment of either goods or time results in zero creation of human capital ($h^C(0, e_0 h_0, \alpha) = h^C(d, 0, \alpha) = 0$), this immediately yields the following:

$$\log(\Delta_\delta h_1) = \log \beta_0^C + \beta_1^C \log \alpha + \beta_2^C \log(e_0 h_0) + \beta_3^C \log d, \quad (1.4)$$

or in levels:

$$h_1 = h_0(1 - \delta_h) + \beta_0^C \alpha^{\beta_1^C} (e_0 h_0)^{\beta_2^C} (d)^{\beta_3^C}. \quad (1.5)$$

Thus, the same ability that helps learning during working life also determines learning ability in college. I assume that ability is more effective in college, meaning $\beta_1^C > 1$. The log-constant term in the above is important, because the role of ability has to nevertheless be rescaled versus that in working life, where the distribution of α is parameterized. Initial human capital, h_0 , is simply assumed to be a linear function of ability, which makes the two perfectly correlated. This is further discussed in the section on parameterization. Because zero goods spending in college results in an ineffective function, the model will endogenously generate a minimum level of spending among college students.

Cost of college As follows from the model of colleges above, the monetary input d depends on the choice variable d but is not the same: here is where we account for institutional aid as well as student grants at the local, state, and federal level. That is why d also depends on ability α and on q , the gross parental wage rate, which is determined by the previous

generation: $q' = wh_t x_t$.⁷ The discussion of the calibration of the model elaborates these points further.

Information structure The agent is uncertain about the next realization of his idiosyncratic earnings state $s_t \in S$. All agents start out from the same state: $s_0 = \bar{s}$. In the next period (when all individuals work) s_1 is drawn from $\Gamma_{s_1}(q)$, where the parental gross wage rate influences the probability of starting out in a good state. This allows the model to capture the importance of parental networks and influence. Thereafter s_t follows a first-order discrete Markov process with transition matrix $\Gamma_s(s_t, s_{t+1})$. The earnings shock $x(s_t)$ combines with his human capital h_t and the wage w of human capital to determine his individual wage rate.

The agent is also generally uncertain about his child's ability $\alpha' \in A$, but gets to know the child's ability right before he makes an inter-vivos transfer. (All choice variables at t^I therefore also depend on α' .) Ability is discrete and drawn from the joint distribution of parents' and children's ability $\Gamma_\alpha(\alpha, \alpha')$.

Taxation A government charges taxes on consumption τ_c , labor income $\tau_n((1 - l_t - e_t)wh_t x(s_t))$, and capital income τ_a . Labor income taxes are non-linear. The government's budget, after consideration of education policies, is balanced by neutral (or wasteful) spending G that does not influence any choices.

Student loans The student loan system mimics the 2003 Stafford loan system as follows. At age 0, college-going students fall into one of two eligibility categories on the basis of their parents' wages at the time they become independent decision makers. If parental wages (q) are not higher than y^* , the student qualifies for subsidized loans up to \underline{b}^s as well as unsubsidized loans up to \underline{b}^u . If parental wages are above y^* , the student can only borrow at the unsubsidized rate up to $\underline{b}^s + \underline{b}^u$. Interest rates r^s and r^u are set exogenously. Interest on subsidized loans is forgiven during the period in which they are paid out. Otherwise, agents cannot borrow at age 0. The model also imposes that those who take out student loans do not save assets at the same time, which is captured by the complementarity constraint $a_1 b_1 \geq 0$. This structure follows and simplifies Abbott, Gallipoli, Meghir, and Violante (2013). After

⁷In reality, policies are heterogeneous across colleges and states, but typically depend on a number of indicators of families' ability to pay for college. Here, the gross parental wage rate is used as a parsimonious proxy. Transitory components would have the potential to make the problem non-convex because parents could adjust their choices to make their children qualify for student aid (which is something policy makers indeed attempt to rule out).

the college-going period, the natural borrowing constraint applies: all loans must be repaid by the end of working life.

After the college-going period, students pay down their debt by a constant amount π every period for m periods. Since pay-down is linear, we can provide an analytical solution for π_t . When $1 \leq t < 1 + m$ and $b_1 < 0$:⁸

$$\pi_t = \begin{cases} -\frac{r^s}{1-(1+r^s)^{-m}}b_1 & \text{if } q \leq y^* \text{ and } -\underline{b}^s \leq b_1 \\ \frac{r^s}{1-(1+r^s)^{-m}}\underline{b}^s - \frac{r^u}{1-(1+r^u)^{-m}}(b_1 + \underline{b}^s)(1+r^u) & \text{if } q \leq y^* \text{ and } b_1 < -\underline{b}^s \\ -\frac{r^u}{1-(1+r^u)^{-m}}b_1(1+r^u) & \text{if } y^* < q \text{ and } b_1 < 0. \end{cases}$$

Otherwise, $\pi_t = 0$. For those who do not enter college, $b_1 = 0$. Finally, in the above $\underline{b}_0 = \underline{b}^s + \underline{b}^u$.

Stationary Equilibrium

The production function takes the following functional form:

$$F(K, H) = K^\theta H^{1-\theta}. \quad (1.6)$$

Here, H denotes the aggregate effective supply of human capital hours. θ is the capital share of total factor income.

Labor, capital, and goods markets are perfectly competitive. We model the economy as closed to labor, and open to capital and goods. This reduces the number of general equilibrium conditions that must be cleared numerically, and is arguably as realistic as assuming an economy that is entirely closed to capital. Additionally, general equilibrium effects through capital formation are by no means a focus of this paper.

Firms borrow capital from households, who receive an international real interest rate r . A share δ of capital is lost to depreciation, which firms reinvest from production. This share is exempt from capital taxation. The openness assumption yields an equilibrium condition relating the capital-labor ratio to the exogenous interest rate, which, together with the income share of labor ties down the marginal product of labor.

For simplicity, all student grants are assumed to be under the control and paid for by the model's government, including institutional aid. The government also issues and collects student debt, pays for public college subsidies, and collects taxes on labor earnings, capital income, and consumption. The government also pays for government expenses, and is

⁸Note that b_1 is a negative number, while \underline{b}^s and \underline{b}^u are positive.

assumed not to hold any government debt or assets other than those mentioned. The government's budget constraint is shown in equation 2.15 below.

Let $x_t^*(\iota_t)$ denote a decision rule given states $\iota_t \in \mathcal{I}_t$ for a generic choice variable x_t . Let \mathbb{I}_t denote a generic subset of the Borel sigma algebra of age-specific state-space \mathcal{I}_t .

Definition 1. A **stationary equilibrium** of the model economy is defined as:

wages w ;

college pricing schedules $\tilde{d}(d, q, \alpha)$;

allocations K, H ;

government spending G ;

net exports NA , net foreign asset position NX ;

decision rules, each $\forall \iota_t \in \mathcal{I}_t$ whenever they are defined for t , for consumption $\{c_t(\iota_t)\}_{t=0}^{T-1}$, leisure $\{l_t(\iota_t)\}_{t=0}^{T-1}$, assets $\{a_{t+1}(\iota_t)\}_{t=0}^{T-1}$, goods $d(\iota_0)$ and time $\{e_t(\iota_t)\}_{t=0}^{T-1}$ investment in human capital, college choice $k(\iota_0)$, student loan borrowing $b(\iota_0)$, and the inter-vivos transfer $v(\iota_{t+1})$;

age-specific measures $\lambda_t(\mathbb{I}_t)$, and the resulting overall measure $\lambda(\mathbb{I})$ on $\mathbb{I} \in \times \mathcal{I}_t$;

such that given international interest rates r , tax functions τ_c, τ_n, τ_a , sets S and A , transition matrices $\Gamma_{s_1}(q), \Gamma_s(s_t, s_{t+1})$, and $\Gamma_\alpha(\hat{\alpha}, \alpha)$, grant schedules $g^I(q, \alpha) \geq 0$, repayment function $\pi_t(b_1)$, as well as the parameters of the model, **the following holds:**

- the decision rules solve the households' problem as described in subsection 1.3.2;
- college pricing schedules solve the colleges' problem; as a result, college pricing schedule take the following form:

$$\tilde{d}(d, q, \alpha) = d - g^I(q, \alpha), \quad (1.7)$$

- the firms make profit maximizing decisions; as a result, their profits are zero and prices of the inputs to production equal their marginal products:

$$r = F_1(K, H) - \delta_a, \quad (1.8)$$

$$w = F_2(K, H); \quad (1.9)$$

- $\lambda_t(\mathbb{I}_t)$ are age-dependent fixed points of the law of motion that is generated by the following:

- the decision rules of the households,
- the laws of motion for assets and human capital,
- the transition matrices of productivity shocks $\Gamma_{s_1}(q)$ and $\Gamma_s(s_t, s_{t+1})$,

- the distribution over the initial states at independence which is consistent with $\Gamma_\alpha(\hat{\alpha}, \alpha)$, parental wealth, and the decisions made by parents on schooling and inter-vivos transfers;

- the market for labor clears:

$$H = \sum_{t=0}^{T-1} \int_{\mathcal{I}_t} (1 - l_t - e_t) x_t h_t d\lambda_t; \quad (1.10)$$

- the market for capital clears:

$$K = \sum_{t=0}^{T-1} \int_{\mathcal{I}_t} a_t d\lambda_t - NA; \quad (1.11)$$

- the balance of payments with respect to the rest of the world holds:

$$rNA = -NX; \quad (1.12)$$

- the market for goods clears (aggregate investment in assets equals depreciation since the equilibrium is stationary):

$$F(K, H) = \sum_{t=0}^{T-1} \int_{\mathcal{I}_t} c_t d\lambda_t + G + \delta_a K + \int_{\mathcal{I}_0} I_{[k=2]} d d\lambda_0 + \int_{\mathcal{I}_0} I_{[k=3]} d^g d\lambda_0 + NX; \quad (1.13)$$

- and the government balances its budget (where the term involving $(d - \tilde{d})$ captures all grants for private colleges, and the term involving $(d^g - \tilde{d}^g)$ captures public college subsidies and grants):

$$\begin{aligned} & G + \int_{\mathcal{I}_t} -I_{[k>1]} b_1 d\lambda_0 + \int_{\mathcal{I}_0} I_{[k=2]} (d - \tilde{d}) d\lambda_0 + \int_{\mathcal{I}_0} I_{[k=3]} (d^g - \tilde{d}^g) d\lambda_0 \\ &= \sum_{t=0}^{T-1} \int_{\mathcal{I}_t} (c_t \tau_c + a_t r \tau_a) d\lambda_t + \sum_{t=0}^{T-1} \int_{\mathcal{I}_t} (n_t w(h_t) x(s_t) \tau_n(\cdot)) d\lambda_t + \sum_{t=1}^m \int_{\mathcal{I}_t} \pi_t d\lambda_t. \end{aligned} \quad (1.14)$$

Finally, I show that the capital market equilibrium condition is satisfied by Walras' law.

- Aggregating individual budget constraints (aggregate investment in assets post-depreciation

is zero since the equilibrium is stationary, inter-vivos transfers net out in aggregate):

$$\begin{aligned}
& \sum_{t=0}^{T-1} \int_{\mathcal{I}_t} c_t d\lambda_t + \sum_{t=0}^{T-1} \int_{\mathcal{I}_t} c_t \tau_c d\lambda_t \\
&= \sum_{t=0}^{T-1} \int_{\mathcal{I}_t} n_t w(h_t) x(s_t) d\lambda_t - \sum_{t=0}^{T-1} \int_{\mathcal{I}_t} (n_t w(h_t) x(s_t) \tau_n(\cdot)) d\lambda_t \\
&\quad - \int_{\mathcal{I}_0} I_{[k=2]} \tilde{d} d\lambda_0 - \int_{\mathcal{I}_0} I_{[k=3]} \tilde{d}^g d\lambda_0 \\
&\quad + \sum_{t=0}^{T-1} \int_{\mathcal{I}_t} a_t r d\lambda_t - \sum_{t=0}^{T-1} \int_{\mathcal{I}_t} a_t r \tau_a d\lambda_t \\
&\quad + \int_{\mathcal{I}_t} -I_{[k>1]} b_1 d\lambda_0 - \sum_{t=1}^m \int_{\mathcal{I}_t} \pi_t d\lambda_t.
\end{aligned} \tag{1.15}$$

- Adding up the government budget constraint and aggregate individual budget constraints:

$$\begin{aligned}
& \sum_{t=0}^{T-1} \int_{\mathcal{I}_t} c_t d\lambda_t + G + \int_{\mathcal{I}_0} I_{[k=2]} d d\lambda_0 + \int_{\mathcal{I}_0} I_{[k=3]} d^g d\lambda_0 \\
&= \sum_{t=0}^{T-1} \int_{\mathcal{I}_t} n_t w(h_t) x(s_t) d\lambda_t + \sum_{t=0}^{T-1} \int_{\mathcal{I}_t} a_t r d\lambda_t.
\end{aligned} \tag{1.16}$$

- Combining this with labor market clearing and profit maximization conditions:

$$\begin{aligned}
& \sum_{t=0}^{T-1} \int_{\mathcal{I}_t} c_t d\lambda_t + G + \int_{\mathcal{I}_0} I_{[k=2]} d d\lambda_0 + \int_{\mathcal{I}_0} I_{[k=3]} d^g d\lambda_0 \\
&= F_2(K, H)H + (F_1(K, H) - \delta_a) \sum_{t=0}^{T-1} \int_{\mathcal{I}_t} a_t d\lambda_t.
\end{aligned} \tag{1.17}$$

- Since $F(K, H) = F_1(K, H)K + F_2(K, H)H$, combining the goods market equilibrium condition with the above yields:

$$\begin{aligned}
& F_1(K, H)K + F_2(K, H)H \\
&= F_2(K, H)H + (F_1(K, H) - \delta_a) \sum_{t=0}^{T-1} \int_{\mathcal{I}_t} a_t d\lambda_t + \delta_a K + NX.
\end{aligned} \tag{1.18}$$

- Combined with balance of payments this yields the capital market equilibrium condition (after rearranging and dividing by $(F_1(K, H) - \delta_a)$):

$$K = \sum_{t=0}^{T-1} \int_{\mathcal{I}_t} a_t d\lambda_t - NA. \tag{1.19}$$

Solution Method

Given the assumptions underlying the above definition of stationary equilibrium, prices have analytical solutions. (The institutional grant component of college pricing schedules is exogenously given.) This leaves the individual's problem to be solved.

The individual problem is a simple life-cycle problem that can be solved by backward induction. At the same time, generations are linked through imperfect altruism. This complicates matters, but not by much: simple rewriting of the problem yields a single recursive equation, which is a relatively standard problem in macroeconomics.

$$V(\alpha, q, a_0) = \max_{\{k(s^0), d(s^0)\}_{s^0}, \{\{\mathbf{z}(s^t)\}_{s^t}\}_{t=j}^{t^I-1}, \{\{\mathbf{z}(s^t)\}_{s^t}\}_{t=t^I+1}^{T-1}, \{\{\mathbf{z}(s^{t^I}, \alpha')\}_{s^t}\}_{\alpha'}, \{v(s^{t^I})\}_{s^{t^I}}} \mathbb{E} \left\{ \sum_{t=0}^{T-1} \beta^t \frac{(c_t^\nu l_t^{1-\nu})^{1-\sigma}}{1-\sigma} + \omega \beta^{t^I} V(\alpha', q', v) \right\}$$

Constraints and transitions are suppressed in the above for parsimony, but are unchanged except that they now depend on $k(s^0)$. The problem can be solved by iterating on an initial guess of V .

The recursive structure combined with individual life cycles increases computational demands, but when solving the problem by iteration on an initial guess the additional burden is reduced by the possibility of introducing *Howard Improvement* steps: one does not need to redo maximization on every iteration, which saves time when the maximization problem is 'large', as is the case here. The full computational procedure is sketched in Appendix 1.8.1.

1.4 Parameterization

I now proceed to discuss the parameterization of the model. The parameterization targets the year 2003 or the closest possible. There reasons for targeting 2003 is data availability: college enrollment in the datasets by Chetty et al. and Hoxby (2016a), ability tests of children in the NLSY dataset, as well as a number of other measurements used in the below all take place close to that year.

The parameter space consists of three parts: Some parameters are estimated outside of the model. These are described in subsection 1.4.1. Some parameters are set directly (either because they have obvious counterparts in reality or because they are readily available in existing literature), and some are set to match moments of the model to their counterparts in the data. These two types of parameters are both described in subsection 1.4.2.

1.4.1 Estimation

A number of important drivers are estimated outside of the model using microeconomic data. These are, in particular, the transmission of ability, the idiosyncratic earnings uncertainty, and the dependency of grants on ability and permanent parental income.

Ability transmission The intergenerational transmission of ability is determined by $\Gamma_\alpha(\hat{\alpha}, \alpha)$. To calibrate this part of the model, I do not choose a functional form. Instead, I directly employ data from the NLSY79 (National Longitudinal Study of Youth '79) and the Children of the NLSY79 datasets, which contain scores on tests taken by mothers and their children. As part of the former study, women aged about 16 to 23 were asked to take an AFQT (Armed Forces Qualification Test) in 1981. They have been tracked since, and their children were also tested using a variety of metrics. This allows to establish a connection between the ability of mothers and their children. The test I use to assess the ability of children is the PIAT Math test, who were between 14 and 16 years old (for the sample we select) when taking the test. I then sort both mothers and children into quintiles on their respective scores, and determine a transition matrix. Figure 1.3 displays the results graphically. Test scores are persistent yet mean-reverting, with a stronger persistence in the tails than in the middle. The overall correlation between mothers and their children's test scores is 0.38.

Because the AFQT score is constructed to generate percentiles, I assume a linear transformation of a discretized standard normal distribution of ability. Specifically, each state is assigned the expected value of an observation in the corresponding quintile of a standard normal distribution. Denoting the discretized standard normal distribution by $\tilde{\alpha}$, and its lowest entry by $\underline{\tilde{\alpha}}$, the distribution of α is formed as follows:

$$\alpha = \tilde{\alpha}\gamma + \rho. \tag{1.20}$$

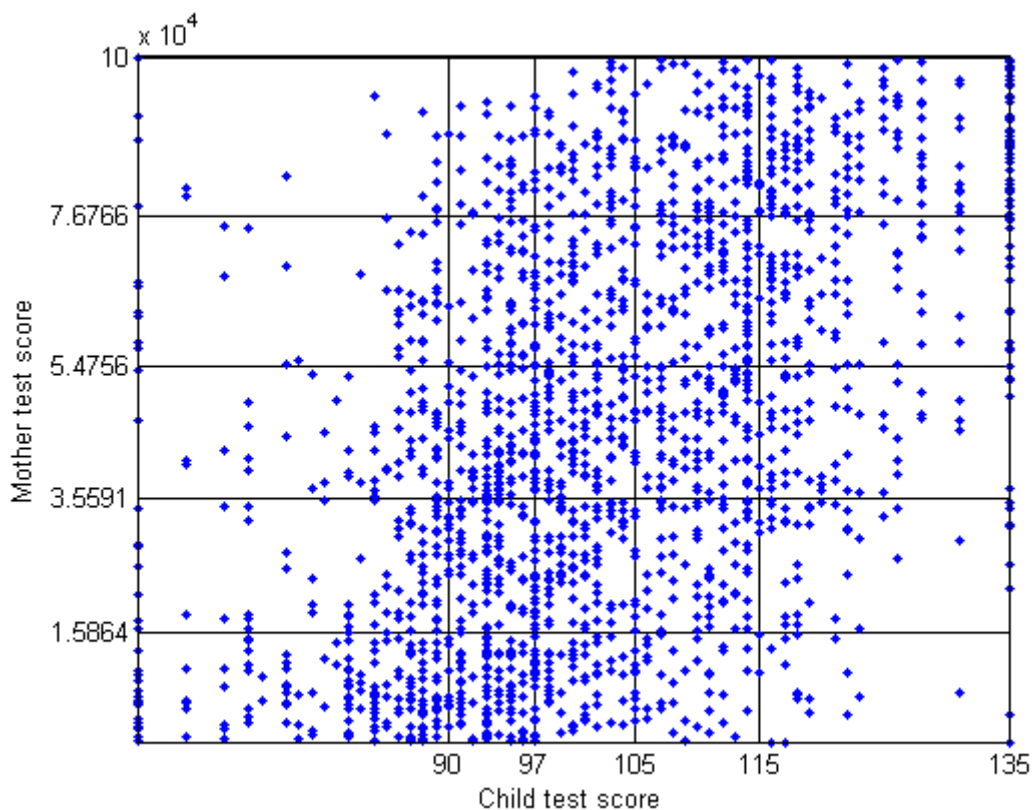
We still need to set the parameters ρ and γ . This is described further below.

Ex ante, there are two issues with the approach. First, these test scores may not actually be a good measure of ability transmission. As we will see, the model (as a positive prediction) produces realistic values of intergenerational persistence by a number of measures. This is perhaps the argument that provides most comfort. In addition, these test scores are commonly used in the literature as measures of ability. In fact, our procedure essentially follows Abbott et al. (2013). It is also worth calling to memory that the procedure is only used to tie down the transition of ability, but not its distribution - which follows from a common functional form assumption.

Second, since ability at the start of adult life is taken as given, one may argue that the Lucas critique applies: policy changes in the model may lead to changes in the behavior of parents

and children at earlier ages. This, in turn, would potentially alter the ability distribution at the start of adult life that we take as given. In that sense, what is called ability here should be interpreted very strictly: it is the transition and distribution of ability as currently measured, i.e. ability at age 18. In practice, policy changes that make learning ability more worthwhile give parents incentives to invest into their children's earlier education. These reactions would likely strengthen the behavioral mechanisms considered in this paper: lifting constraints on educational investment in adult life will make earlier investments more valuable as well (assuming different stages of education are complements). From that perspective, the effect sizes reported in this paper will be conservative.

Figure 1.3: Ability transition



Grid lines show quantile bounds for each of the tests. The density of observations in each rectangle of the grid illuminates persistence in test scores across generations.

Earnings uncertainty For most of the life cycle, the model setup of this paper restricts idiosyncratic earnings uncertainty to be of a first-order Markov form, so that only one state is required to track the idiosyncratic component of earnings. This process is ideally calibrated based on an empirical study of hourly wages that allows for significant heterogeneity in the systemic component of wage profiles. As Guvenen, Kuruscu, and Ozkan (2014) note, the

closest such study is by Haider (2001). Two complications now arise: that paper uses an ARMA model for log wages, which would take an additional state variable to track the moving average of wages, and its estimates are based on yearly data while the calibration period in this paper is four years. These issues are resolved as follows: the ARMA process estimated by Haider (2001) is simulated, after which every four simulated periods are summed to one, and an AR(1) process is estimated on the resulting series using maximum likelihood. Taking this approach, we use both the best possible measurement of the idiosyncratic component of wages, and the best possible approximation of that process in the context of our model. The estimates of the autoregressive coefficient and error term variance are then used to create a discrete and symmetric first-order Markov process with two states, which has the same persistence and unconditional variance as the estimated AR(1) model. The final $\Gamma_s(s_t, s_{t+1})$ and $x(s_t)$ are shown in Table 1.2. The initial state of s in the model is fixed and denoted \bar{s} , and is set to the lower of the two states.

Table 1.2: Idiosyncratic earnings process

		To	
		1	2
From	1	0.72	0.28
	2	0.28	0.72
Value		0.72	1.28

The careful approach above is chosen because these shocks are an important element of this paper. The yearly persistence implied by the the four-year probability of remaining in the same state is 0.92. This is in line with other findings in the literature. For example, Storesletten, Telmer, and Yaron (2004) report a yearly autocorrelation of 0.95.

Student grants and college subsidies Appendix 1.8.2 provides an overview of the landscape of US education policy around 2003. In part, student aid consisted of student loans and subsidies to public colleges, which are modeled explicitly in this paper and parameterized further below. For the remained, a plethora of student grants (from federal, state, and local governments, as well as tuition discounts based on family income and merit) create a wedge between individual costs for college (\tilde{d}) and actual investment in human capital (d). I now lay out the mapping between these two variables, and then parameterize it by estimation from data.

First, let us call the sticker prices observed in the data (a college’s headline figure for tuition and fees) s^D (where the superscript D will refer to data). Next, let us relate total aid g^D (from colleges and all levels of government) to sticker prices to capture general subsidies, and

also to family income and human capital. These latter two capture need- and merit-based aid. I consider a linear relationship as follows:

$$g^D = \beta_0 + \beta_1 q^D + \beta_2 s^D + \beta_3 \alpha^D. \quad (1.21)$$

Given data on grants, we can estimate the parameters in this equation.

Because sticker prices are paid either through private expenditures or from aid (which we have defined broadly), we have $s^D = g^D + \tilde{d}^D$. In addition, competitive pricing schedules guarantee that $d^D = g^D + \tilde{d}^D$, so that $s^D = d^D$. Concluding all of this, investment in college is determined as follows:

$$d^D(\tilde{d}^D, q^D, \alpha^D) = \frac{1}{(1 - \beta_2)} [\beta_0 + \beta_1 q^D + \tilde{d}^D + \beta_3 \alpha^D]. \quad (1.22)$$

We still need to connect the data variables to those in the model. Here, there are two issues at play. First, the numeraire in the model is different from the numeraire in the data. Second, the unit of measurement for human capital will be different. I resolve this by rewriting equation 1.22 as follows:

$$\frac{d^D}{\bar{y}^D} = \frac{\beta_0}{(1 - \beta_2)\bar{y}^D} + \frac{\beta_1}{(1 - \beta_2)} \frac{q^D}{\bar{y}^D} + \frac{1}{(1 - \beta_2)} \frac{\tilde{d}^D}{\bar{y}^D} + \frac{\beta_3}{(1 - \beta_2)\bar{y}^D} \bar{\alpha}^D + \frac{\beta_3 \sigma_\alpha^D}{(1 - \beta_2)\bar{y}^D} \frac{\alpha^D - \bar{\alpha}^D}{\sigma_\alpha^D}. \quad (1.23)$$

Here, \bar{y}^D are average earnings as measured in the data. $\bar{\alpha}^D$ and σ_α^D are also assumed measurable in the data, and represent the mean and standard deviation of α^D . Now, note that this is an equation relating normalized instructional expenditure $\frac{d^D}{\bar{y}^D}$ to normalized parental income $\frac{q^D}{\bar{y}^D}$, normalized personal education expenditure $\frac{\tilde{d}^D}{\bar{y}^D}$, and normalized ability $\tilde{\alpha}^D = \frac{\alpha^D - \bar{\alpha}^D}{\sigma_\alpha^D}$. All of these terms have clear model counterparts, while the coefficients are measurable in the data.⁹

Rewriting for the model counterpart of equation 1.23, we get (with the superscript M referring to model variables):

$$d^M = a_0 + a_1 q^M + a_2 \tilde{d}^M + a_3 \tilde{\alpha}^M. \quad (1.24)$$

Here, $a_0 = \frac{\beta_0}{(1 - \beta_2)} \frac{\bar{y}^M}{\bar{y}^D} + \frac{\beta_3}{(1 - \beta_2)} \frac{\bar{y}^M}{\bar{y}^D} \bar{\alpha}^D$, $a_1 = \frac{\beta_1 \bar{n}}{(1 - \beta_2)}$, $a_2 = \frac{1}{(1 - \beta_2)}$, and $a_3 = \frac{\beta_3 \sigma_\alpha^D}{(1 - \beta_2)} \frac{\bar{y}^M}{\bar{y}^D}$. All inputs underlying these terms can be estimated from data.

Next, we turn to measurement. The National Postsecondary Student Aid Study (NPSAS) by the NCES for the year 1995-1996 links surveys of student finances to characteristics of the colleges they are enrolled in. In this dataset we find total aid received from all sources

⁹The counterpart to q^D , gross parental income, is q^M , a gross wage rate. Thus, we use $q^M \bar{n}$ as the relevant counterpart to turn the model wage rate into model earnings. We set \bar{n} to 0.35, based on a daily time endowment of 16 hours (for each of 7 days) and a reported weekly 39.53 hours of total market work in 2003 Aguiar and Hurst (2007).

(except Stafford and PLUS loans), tuition and fees (before any aid), gross parental income, as well as SAT scores (combined scores) which function as a proxy for human capital. I use these data to estimate equation 1.21 by Ordinary Least Squares, restricting the sample to 4-year colleges. The regression is done separately for private and public colleges. Observations containing zeros are excluded, except for grants. The regression is weighted by the NCES's full sample weights.

Table 2.3 contains the estimates, as well as a measure of the explanatory power of the linear model and the number of observations used. Finally, the resulting parameters of equation 1.24, which are directly fed into the model, are displayed as well. From the NPSAS we have that $\sigma_\alpha^D = 226.1$ and $\bar{\alpha}^D = 930.0$ when assuming a normal distribution on the SAT score data (calculated from percentile data), which is also the assumption in the model. \bar{y}^D is \$31,141 in 1995 USD according to the OECD.

Table 1.3: Regression results (standard errors in brackets)

	(1.21) g^D			(1.24) d^M	
	Public	Private		Public	Private
Constant	1127.10 (331.59)	376.88 (815.61)	Constant: a_0/y^M	0.05	0.11
q^D	-0.01 (0.01)	-0.02 (0.02)	$q^M : a_1$	-0.01	-0.02
s^D	0.15 (0.02)	0.24 (0.04)	$\tilde{d}^M : a_2$	1.18	1.32
α^D	0.19 (0.38)	2.39 (0.72)	$\tilde{\alpha}^M : a_3/y^M$	0.00	0.02
R^2	0.09	0.12			
Observations	~5,600	~4,800			

The regression results show that grants for private colleges depend more on merit and need compared to those for public ones, but that the latter have a larger constant component. This conclusion regarding the intercept changes slightly when translating the regression results to model parameters in the right half of the table. Because sticker prices are higher in private colleges, grants tend to be higher as well even when disregarding merit and need. In both types of colleges, not spending anything results in a positive grant when parental income is zero. Also, spending more results in more investment (since that is a one-to-one relationship) but also in more need and thus more grants, making the coefficient of spending larger than one. Negative grants could technically occur in this linear relationship for some combinations of inputs, but do not actually occur in the calibrated model.

1.4.2 Moment Matching

The below describes the moments used, together with the parameters that they are informative of. This subsection ends with an overview.

Life-cycle The model period is set to four years. Model ages are set as close as possible to their counterparts in reality: working life starts at age 18 ($t = 0$), retirement at 66 ($T = 12$). Child birth occurs at age 28, which is the average age of mothers at child birth¹⁰, so that children start their working life when the parent is aged 46. Inter-vivos transfers are made during the period before that ($t^I = 6$).

Production We use values for discounting (β , yearly value 0.987) and depreciation (δ_a , yearly value 0.012) that are standard in the literature. We adjust these values for our model period. The international interest rate (r) is set such that the post-depreciation yearly rate r is 1%. This results in an interest rate slightly below that of an equivalent closed complete markets economy ($\frac{1}{\beta} - 1$). θ is set equal to the capital share of total factor income in the data (0.33).¹¹

Preferences ν and σ are set to match average hours worked and the elasticity of intertemporal substitution. The former is taken to be 35%, based on a daily time endowment of 16 hours and a reported weekly 39.53 hours of total market work in 2003 Aguiar and Hurst (2007). For the latter we rely on a meta-study by Havránek (2015), who finds that the literature's best estimate for this elasticity is 0.3-0.4 after correcting for publication bias. I use the midpoint of that range.

Inter-vivos transfers Abbott et al. (2013) do extensive empirical work on inter-vivos transfers using survey data from the NLSY97.¹² They estimate average total inter-vivos transfers between age 16 and 22 to be \$30,566 in 2000 dollars (79% of the 2000 average wage, or 20% of 4 years of average wage when accounting for the model period), and we set ω to match this figure with our one-off inter-vivos transfer.

¹⁰Calculated from 2010 data provided by the Center for Disease Control and Prevention (CDC).

¹¹Data are available from the OECD for 2003.

¹²The NLSY97 surveys a nationally representative sample of individuals in much the same manner as the NLSY79, starting in 1997. Participants were aged 12 to 16 when they first participated.

Human capital The initial distribution of human capital (h_0) is assumed to be a linear transformation of the distribution of ability, and thereby perfectly correlated with ability. Here, the paper essentially takes the view that it is ability to learn that is, together with actual knowledge, built earlier in life. Once the child matures, the two are then separate entities: underinvestment can lead to a level of knowledge that is low versus learning ability, and vice versa. If we were to let go of the link at an earlier age, catch-up effects might occur where an undertrained but able child, given the same educational, outperforms peers who are more knowledgeable to begin with. While this may certainly occur in practice, we choose to ignore the effect here: First, the empirical literature points in another direction, suggesting that there are strong complementarities between early and later education. Indeed it seems that the purpose of training in early childhood is in large part 'learning to learn' what is taught in tertiary education and at work. Second, related papers that separate ability and initial human capital early in life, such as Huggett, Ventura, and Yaron (2011), find the two to be strongly correlated. Other papers have therefore proceeded in the same way as I do, notably Guvenen, Kuruscu, and Ozkan (2014).

Quantities of human capital are yet to be normalized, which is done as follows:

$$h_0 = h_{norm} + (\tilde{\alpha} - \underline{\alpha})\psi. \quad (1.25)$$

Thus, the lowest level of initial human capital in the economy is normalized to h_{norm} . The resulting normal distribution (approximate due to discretization) has mean $(h_{norm} - \underline{\alpha}\psi)$ and standard deviation ψ . These results are used to implement equation 1.24.

Summing up, the parameters γ, ρ (from equation 1.20) regulate the distribution of ability, the parameter ψ (from equation 1.25) regulates the distribution of initial human capital (while h_{norm} can be set to any computationally convenient value), and β^w and δ_h (from equation 1.3) regulate the build-up of human capital while at work. I set all of these parameters to capture features of the distribution of age-earnings profiles.

Huggett, Ventura, and Yaron (2011) do empirical work to establish the distribution of patterns of life-cycle earnings, taking into account time fixed effects. These data are displayed in Figure 1.10 (in a later section of the paper). The sample consists of men who are attached to the labor force. They show: (i) that earnings increase and then decrease over the life cycle, (ii) how large this movement is versus what is given at the beginning of the cycle, (iii) that inequality grows with age, and (iv) how much inequality there is in the system overall. The model equivalents of these patterns are driven by the distributional parameters above. I take the following moments from the data that capture these patterns:

1. Average earnings at age 32 over average earnings at age 24. (1.37)

2. Average earnings at age 48 over average earnings at age 24. (1.57)
3. Average earnings at age 60 over average earnings at age 24. (1.32)
4. The variance of log earnings at age 32. (0.34)
5. The variance of log earnings at age 48. (0.42)

College effectiveness The effectiveness of college, together with the life-cycle of earnings, is informative of the extent to which human capital is determined before college. The constant elasticity functional form in equation 1.5 leaves the following parameters to be determined: β_0^C , β_1^C , β_2^C , and β_3^C .

Key to some of the questions this work is after is the relative importance of financial resources in the college production function of human capital. Hoxby (2016a) identifies the effectiveness of money across the distribution of colleges in a setting where financial investment is approximately exogenous, meaning that ability is controlled for. I target these results, which are described further below.

β_0^C determines how effective ability is in college versus at work, so that the share of the population that decides to go to college is informative. Using data from Chetty et al. (2017) which is on the relevant cohorts, I find that 75% of individuals in the relevant cohort enroll in some sort of college. This is the relevant empirical counterpart for the model.

It is generally worth noting at this point that the paper takes a broad view of human capital: human capital is continuous, and I do not explicitly deal with dropouts, 2-year colleges, professional degrees, etcetera.

Combining all this, average inputs of time and money then imply the remaining parameters:

1. Data on time use by students are hard to come by, and do not generally paint a consistent picture. Perhaps most important here is to capture the ability of students finance their education by work time. I use the 2003 American Time Use Survey, and restrict the sample to those enrolled in college and spending at least some time attending class. I then calculate how much time these students spend on education (including education-related travel) versus work (including work-related travel), aggregating individuals using ‘ATUS final weights’. The ratio of the former category versus the latter is 2.02: active students spend about twice as much time studying as they spend working. I then halve this ratio twice: once to account for time during which colleges are out

of session¹³, and once to account for time actually spent in college during the 4-year period¹⁴.

2. To tie down spending on education, the calibration targets the share of GDP spent on tertiary education from private sources. Because private spending in our model is very narrowly defined as direct spending by households, we take the NIPA account on private spending on higher education for 2003 as the counterpart in the data, which is 0.86% of the 2003 average wage.

Hoxby (2015, 2016a, 2016b) measures causal returns of a marginal dollar investment upon college entry from discounted lifetime income. Her method is as follows: combining administrative data on incomes, clearinghouse data on college applications, and data on college expenditures, she compares students who are ‘on the bubble’ of getting admitted to a college. Student SAT scores help to identify students who are close to being admitted or rejected. The assumption that identifies causal effects is the following: For students whose credentials are close to the typical cut-off, admission can be thought of as a random event. Paired comparison methods then establish the extra monetary investment caused by admission, and the subsequent returns to that investment. In doing so, the least selective college is normalized to add zero value. Results show a marginal dollar return of around 3.5 after discounting for colleges that are at least somewhat selective, and these returns increase slightly in college selectivity. The results are relevant to college entry, so that college dropouts, 2-year or 4-year colleges, and all other such issues, are averaged out.

Hoxby’s results have a model equivalent: I simulate the effect of an exogenous extra dollar investment in college on discounted lifetime income (using the same discounting method as Hoxby). The average of the resulting returns is a model moment that can be set to match the typical return reported in Hoxby (2016b).

Public college Two choices are required regarding the representative public college. What is the cost of attending before any grants (the sticker price), and by how much does the government directly subsidize the college? According to Johnson (2014), the subsidy rate for an average public college is about 53%, which leaves some \$5,640 a year (or 10.48% of average earnings) of an average \$12,000 in spending per student per year to be paid for by students and grants (data for 2011-2012). Therefore, \$5,640 a year is the average sticker price.

¹³The assumption here is that time worked stays constant over the year, and time spent studying goes to zero during half the year. The exercise remains approximate due to the lack of appropriate aggregate data.

¹⁴This is to correct for drop-outs from college, 2-year-colleges, etcetera. Again, the exercise remains approximate due to the lack of appropriate aggregate data.

Student loans I follow the structure of Abbott, Gallipoli, Meghir, and Violante (2013) to model the student loan system around 2003, but with some simplifications. The main simplification is that I do not model private student loans. As detailed in Appendix 1.8.2 and in Abbott, Gallipoli, Meghir, and Violante (2013), these were a small source of financing, and mostly available to students whose parents were sufficiently credit-worthy. Indeed, including them would only strengthen the conclusions of this paper (but it would also come at added computational cost).

The parameters y^* , \underline{b}^s , \underline{b}^u , r^s , and r^u are informed by the following moments:

1. Stafford aid was 0.33% of GDP or 0.31% of the average wage (College Board, 2013).
2. Subsidized Stafford aid was 54.54% of total Stafford aid (College Board, 2013).
3. The subsidized loan limit over the unsubsidized loan limit, which was 0.95.¹⁵
4. Interest rates for either type of student loan was around 4% (or 17% on a 4-year basis) in 2003.¹⁶

The repayment period length m is set to 20 years. While the initial repayment period has typically been 10 years, this is easily extended in practice.

The above moments with regards to the Stafford loan system are chosen to accurately represent the generosity of the program overall, as well as to specific family income groups. Cumulative loan limits for the two types of Stafford loans exist, and we use these to tie down the relative generosity of the two programs in terms of available funds. However, whether students can borrow up to these limits depends on a number of other factors (for example their class level, dependency status, cost of attendance, and financial need), so that we instead focus on matching overall amounts of borrowing. Costs of student loans (interest rates) are taken from the data, ensuring an accurate representation of that aspect.

Tax policies Guvenen, Kuruscu, and Ozkan (2014) collect data on US earnings taxes at different income levels for the year 2003 from the OECD. Using these data, I directly estimate the two parameters of the much-used tax function described in Heathcote, Storesletten, and Violante (2017). All of that results in the function below, where \bar{y} are the average United

¹⁵According to FinAid (2016), the aggregate subsidized loan limit in 2003 was \$17,125, which was 43.16% of GDP per capita at the time, while the aggregate unsubsidized loan limit in 2003 was \$18,000, which was 45.37% of GDP per capita at the time.

¹⁶According to FinAid (2018), interest rates for Subsidized and Unsubsidized Stafford loans were the same in the early 2000s, the rate being 4.06% in 2002-2003 and 3.42% in 2003-2004.

States earnings (and the same parameter that is used in the implementation of equation 1.24).¹⁷

$$\tau_n(\cdot, \bar{y}) = 1 - \frac{1}{1.3434} \left(\frac{n_t w h x_t}{\bar{y}} \right)^{-0.11867}.$$

I take the consumption and capital income tax rates from McDaniel (2007): $\tau_c = 0.075$, and $\tau_a = 0.232$ for 2003. Government expenditures G are calibrated to clear the government budget.

Overview Table 1.4 provides an overview of parameters set outside of the model and their values. Table 2.4 lists parameters that were set to match moments: it displays the final parameter values, together with the moments as measured in the model and in the data. Percentages refer to either average wage or average wage per capita.

Table 1.4: Parameters set outside of the model

	Value	Moment
β	0.949	Discount rate
δ_a	0.047	Asset depreciation rate
r	0.088	Pre-depreciation real interest rate
θ	0.314	Capital share of income
r^s	0.170	Subsidized Stafford loan rate
r^u	0.170	Unsubsidized Stafford loan rate

Few of the parameters in Table 2.4 have a natural interpretation. The value for ω suggests that parents count their children's value function for a fifth of their own at the age where the children mature. Depreciation of human capital is 12% during a four-year period. The elasticity of human capital growth in college is largest in ability, followed by money and time invested.

¹⁷I simply apply this formula to periodic model incomes, normalized by average wage. This would be equivalent if incomes were indeed constant during the model period. For the purposes of this paper, we consider it a good enough approximation.

Table 1.5: Parameters set to match moments

Parameter	Value	Moment	Model	Data
σ	3.32	Elasticity of intertemporal substitution	0.35	0.35
ν	0.80	Average labor supply	0.41	0.39
ω	0.20	Average inter-vivos transfer	18%	20%
γ	0.90	Variance of log earnings: age 32	0.34	0.34
ρ	1.90	Average earnings: age 48 vs. age 24	1.58	1.57
β^W	0.10	Average earnings: age 32 vs. age 24	1.39	1.37
δ_h	0.12	Average earnings: age 60 vs. age 24	1.45	1.32
ψ	0.40	Variance of log earnings: age 48	0.48	0.42
β_0^C	0.70	Share with some college	76.11%	74.65%
β_1^C	1.58	Share of average wage spent on tuition	0.86%	0.83%
β_2^C	0.11	Time spent on education versus work	0.46	0.51
β_3^C	0.66	Average causal dollar return	3.18	3.50
$\tilde{d}^g(d^g, 0, 0)$	0.73	Average sticker price tuition	10.48%	10.48%
d^g	1.26	Average subsidization rate	0.53%	0.53%
y^*	25.00	Subsidized versus Unsubsidized Stafford aid	50.16%	54.54%
\underline{b}^s	0.27	Subsidized versus unsubsidized loan limit	0.95	0.95
\underline{b}^u	0.28	Overall Stafford aid	0.36%	0.31%

1.5 Implications

This section considers positive implications of the model that have not been targeted in parameterizing the model.

1.5.1 Intergenerational Mobility

No measure of intergenerational persistence of economic outcomes has been targeted in the model's parameterization. This subsection compares model predictions against actual measurements of persistence. The success of the model, summarized below, provides confidence in the ability transition matrix, that was based on test scores.

Table 1.6 contains several measures of intergenerational persistence, first for the model and then for the data. A range of estimates of the IGE exists in the literature. In their reviews of the literature, Lee and Seshadri (2014) and Landersø and Heckman (2017) respectively arrive at ranges of 0.4–0.6 and 0.3–0.5. The range of estimates is large due to differences in sample selection and treatment: one can restrict the ages at which earnings are measured, the labor attachment of individuals, their gender, etcetera. I report the model IGE using the entire population, and then again controlling for age. Rank correlation measures are typically a bit

Table 1.6: Measures of Intergenerational Mobility

	Model	Data
Intergenerational earnings elasticity (IGE)	0.34	0.3–0.6
IGE with controls for age	0.31	
Intergenerational rank correlation of earnings	0.30	0.34
Correlation in educational attainment	0.21	0.11–0.45

lower. One such measure, by Chetty, Hendren, Kline, and Saez (2014) comes out at 0.34, in line with our model (again using the entire model population). Finally, the intergenerational correlation of educational attainment can also be measured. For the model, we do so simply using indicators of college entry. As expected, the outcome is a bit below the other measures and well within the range reported in the literature (see Mulligan (1999, Table 1)).

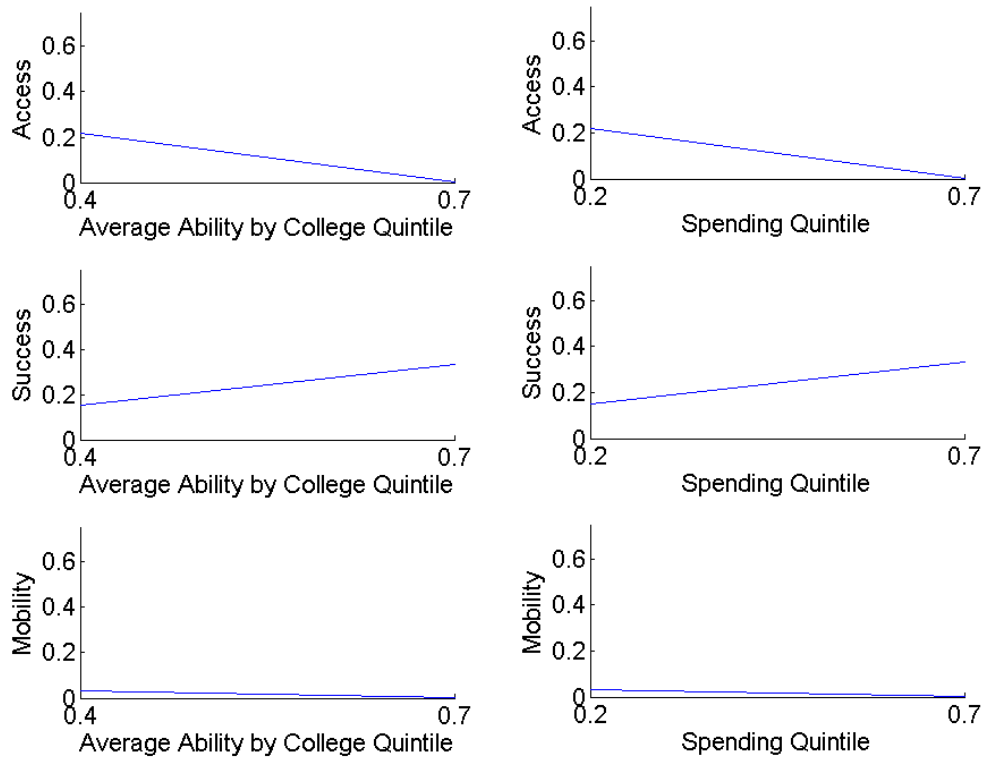
1.5.2 Heterogeneity of Intergenerational Mobility

Figure 1.4 reports the model equivalents to the measures used in Chetty et al. (2017), which were displayed in Figure 1.2. Instead of grouping students by the college they go to, I form quintiles of students by college spending (since spending is what identifies a college in the model).¹⁸ I then measure average ability by college quintile (where individual abilities have been normalized to have mean zero and a standard deviation of one.). The patterns we see are qualitatively and quantitatively in line with their empirical counterparts. The share of low income students (defined as a family income in the bottom 20%) amongst those that go to college (‘access’) falls as the average ability of students in a spending quintile. At the same time, the likelihood that a student from a low income family reaches the top 20% of the income distribution (‘success’) rises in the average ability of students. The product of the two, the share of college students that go from the bottom to the top of the distribution (‘mobility’), is flat across ability. The same patterns hold for the spending quintiles themselves, where spending has been normalized by the average labor earnings in a model period.

Chetty et al. (2017) also discusses how the intergenerational earnings elasticity changes when one controls for the college a student enters. In doing so, they reduce the sample to those who enter any college. Intuitively, if college choice is a perfect measure of human capital, and human capital is all there is to earnings, one would expect the IGE to be reduced to zero. If financial constraints make able students enter worse colleges, then in a given college the children of the poor might even out-earn the children of the rich, making the IGE negative. Chetty et al. report a national IGE (based on rank-rank regressions) of 0.29, which is reduced

¹⁸There are few distinguishable quintiles due to bunching in the public college.

Figure 1.4: Access, Success, and Mobility in the Model



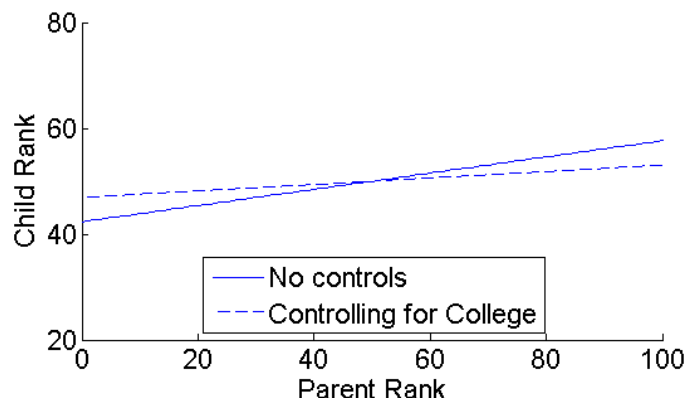
by two-thirds to 0.1 when including college fixed effects. (See Figure 1.1 and the discussion there.) Should this result be taken to indicate that human capital cannot explain all of the earnings persistence? The model suggests another explanation.

Figure 1.5 illustrates a similar procedure, but on model-generated data. College choices in the model are very granular. So instead of using college fixed effects, educational investment is included as a control. Again, one would expect to find a flat or even declining line, since in the model all persistence is due to human capital. Interestingly, the coefficient also remains positive in the case with controls, just as in the data.

Key to understanding this is the following: college spending and ability are not perfectly assortative, even in the absence of constraints. While college spending is more effective for the more able, their optimal level of investment is nevertheless not necessarily higher. This is because time and money can be transformed through wages in the model, so that optimal investment level also depend on demands for leisure time. Substitution effects dominate income effects in the model overall, so that the more able earn and work more later in life. At the same time, it is optimal for them to enjoy more leisure relatively early in life, as their expected wages grow steeply. These increased demands for leisure in college can undo the

higher effectiveness of spending. In short, smarter students sometimes study less and enjoy more leisure, even when they can afford to go to a more expensive college. As a result of the imperfect assortativeness between ability and investment, controlling for colleges does not actually control for human capital. The initial premise, that one should expect the slope to be zero when human capital alone explains persistence, is false: it can be positive, even when constraints are present.

Figure 1.5: IGEs with and without controls



Graphs are created by fitting a straight line with slope equal to the IGE estimate through the point (50,50).

Landersø and Heckman (2017) estimate the IGE non-linearly, and find that persistence is larger for those with higher income parents. A simple quadratic regression on model-generated log earnings indeed produces a convex relationship between the earnings of two generations: earnings persistence is stronger when parents have high earnings. Such findings are entirely in line with one of the main messages of this paper: at the top of the earnings distribution, human capital is more important than other components of earnings. At the same time, human capital itself is quite persistent. Thus, we quite naturally find higher persistence at the top of the distribution.

1.5.3 College Entry and Heterogeneity

Lochner and Monge-Naranjo (2011b) shows the gradient of college enrollment by measured ability and family income empirically, based on NLSY97 data for the early 2000s. Enrollment rises strongly in ability, but also in family income. Figure 1.6 displays the model equivalent.¹⁹ Because ability is perfectly observable, the gradient in ability is much stronger than in the data. As a result, the gradient in family earnings is not visible by overall enrollment. However, as the results on public college enrollment and spending below will indicate, constrainedness

¹⁹It is worth noting that family earnings here are not equivalent to q from before, due to labor supply and a difference in timing of measurement.

(which I directly observe in the model) does depend on family income.²⁰

Figure 1.6: College Entry by Ability and Family Income

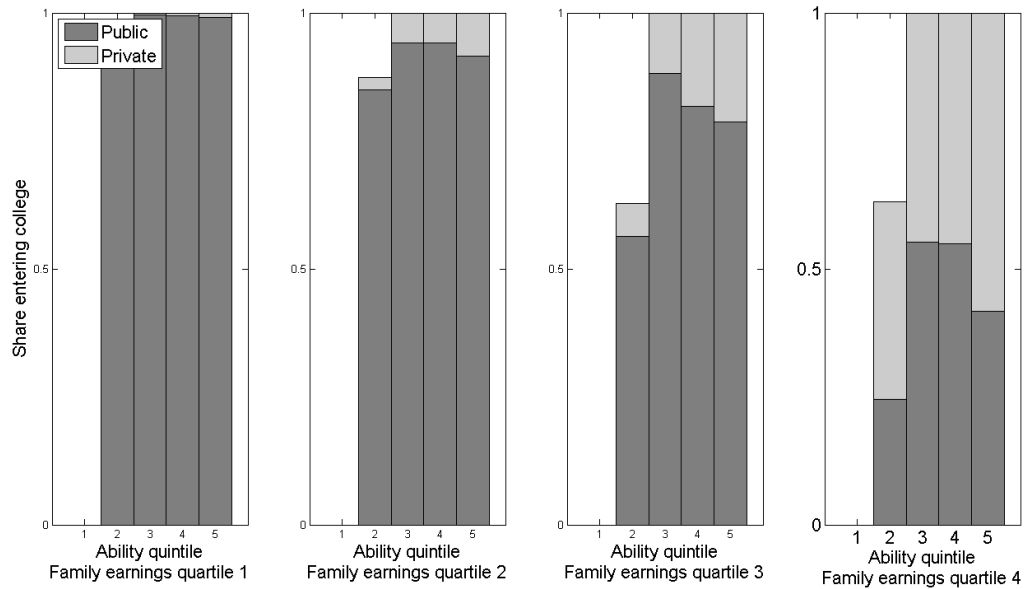
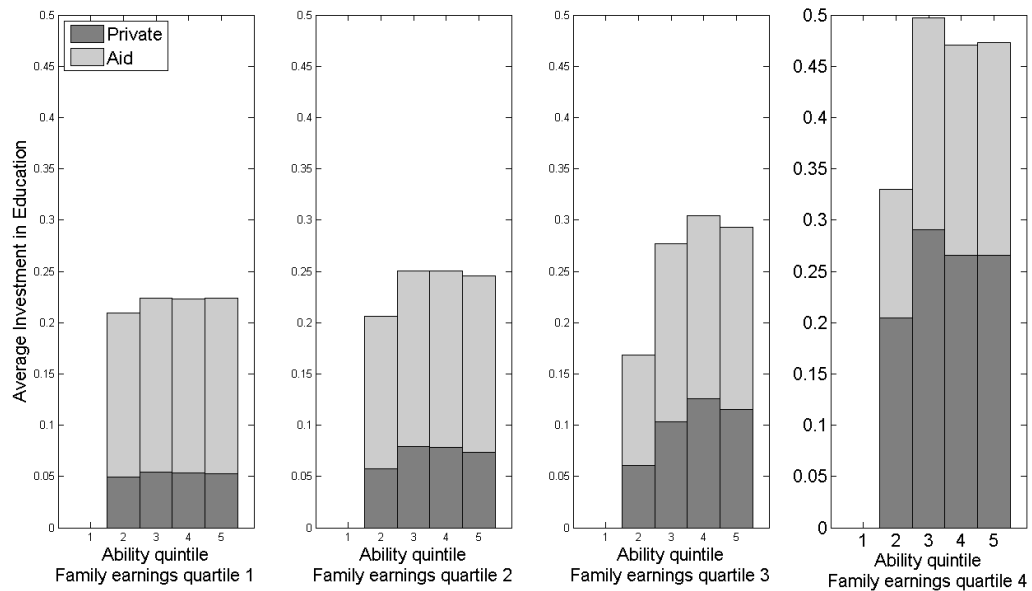


Figure 1.7: Average Investment in Education by Ability and Family Income



²⁰The model generates a surprising enrollment pattern for those in the second ability quintile. Members of this quintile that do not enroll into college are exclusively agents who have not received any transfers from their parents. At the same time, their access to grants reflects their parents' income position, leading to a negative gradient in family income.

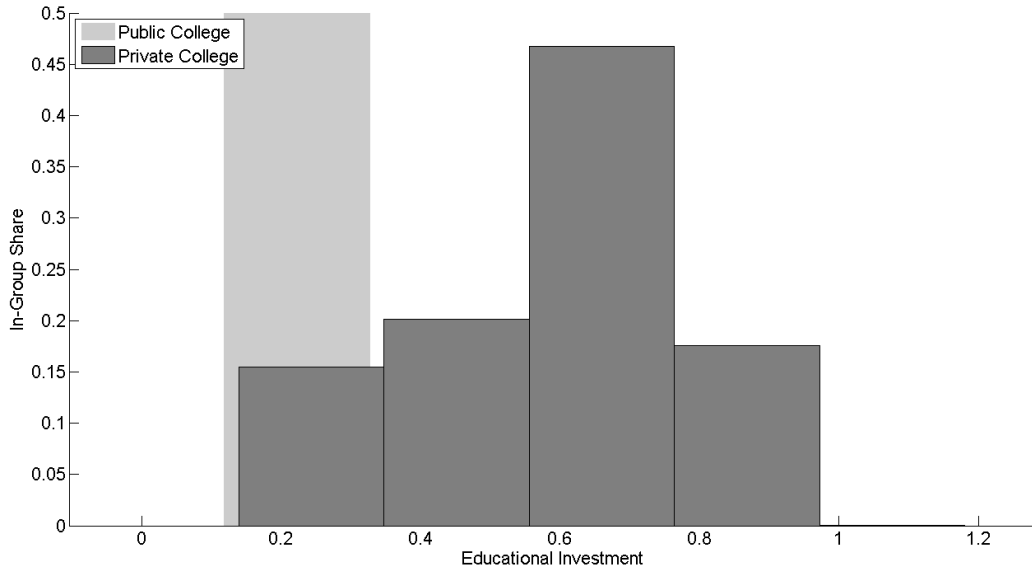
Figure 1.6 also splits entry between public and private colleges. In the data (e.g. the cohorts analyzed in Chetty et al. (2017)) almost 80% of students go to ‘public’ colleges, although this category does also include 2-year private not-for-profit institutions. As we would expect, children of lower income families tend to go to the public colleges. A similar pattern is visible in Figure 1.7, which shows goods invested in education by the same split. Investment is expressed in terms of 2003 average wages, which is \$168,780 for the four year period. Investment is split in private spending (including spending financed by loans) and aid, which consists of both grants and subsidies in the case of public colleges. Spending heterogeneity is mostly driven by family resources, and less so by ability. Aid is remarkably stable over both gradients. This is due to three factors. First, there is only one representative public college, so that subsidies are the same for all that go there. Second, grants for private colleges are much larger, bringing aid for the students that go there into the same range. Third, for both types of colleges, aid schedules are dominated by the intercept, meaning they are roughly the same for all their students.

It is fairly obvious that a pattern of increasing college entry in ability could be achieved by introducing ‘preference shocks’ that depend on parental background, as is typical in some parts of the literature. I do not pursue this for three reasons: First, it makes the model less parsimonious. Second, it is not obvious to what extent constraints, measurement error (since SAT scores and the like are an imperfect measure of ability), and preferences are each responsible for the pattern observed in the data. Third and most importantly, the results presented here bear out one of the key results of the paper: even when there are no constraints to overall college enrollment (the choice to go to college or not, which I call the *extensive* margin of college choice), students may still be significantly constrained in their choice of a particular college (which I call the *intensive* margin of college choice). After all, investment in education still depends on family earnings.

For completeness, Figure 1.8 displays the distribution of private colleges by their total educational investment. The investment level of the representative public college is highlighted as well (normalized by the average labor earnings in a model period). Private colleges invest more in students than public ones. Their distribution is increasing at lower levels of spending, which is in part due to students at the lower end that crowd into the public college.

Finally, returns to education are also heterogeneous by college. The approximate average of the marginal per-dollar returns reported in Hoxby (2016b) was a target in the parameterization of the model. That paper also shows some heterogeneity of returns to college, with the per-dollar return growing in the average SAT scores of incoming students. The same pattern is present in model-generated data: marginal per-dollar returns grow in students’ ability.

Figure 1.8: Distribution of Colleges

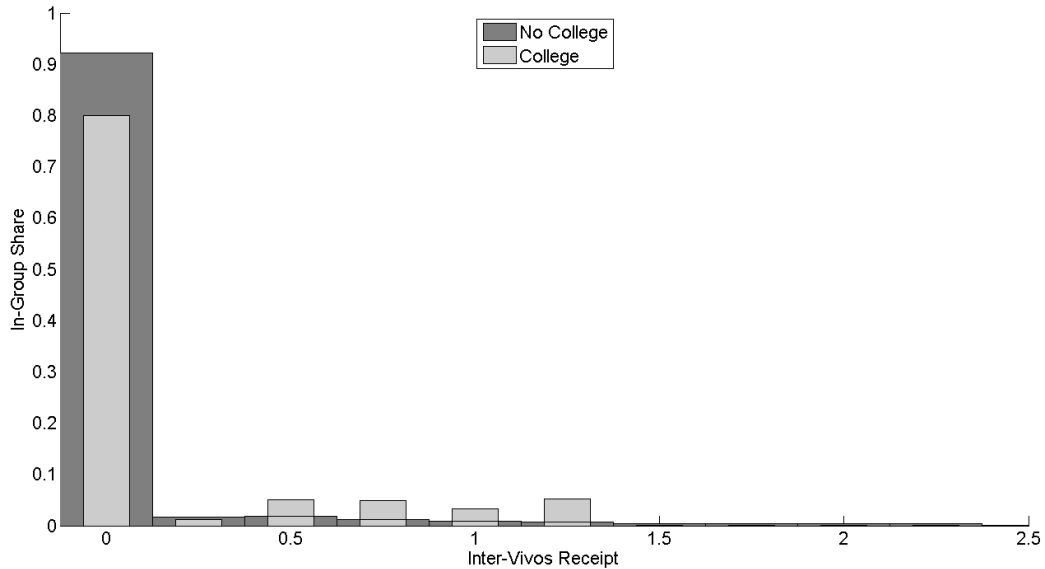


1.5.4 Paying for College

Within the model, constrainedness can be measured directly. Even in the highest family earnings quartile, a significant share of students is constrained in their choices. Of course, the regression methodology of this paper discards much heterogeneity in grants that may be relevant in practice. The overall impact of this on constrainedness is unclear. In any case, the patterns generated by the model points to an interesting nuance. Constraints seem to keep few from enrolling in at least some college. (This is in some part due to the presence of the public option.) At the same time, almost all spend less on college than would be optimal, and returns to college remain large (as in the data). Research on financial constraints should keep this in mind: it is not just the exogenous (enrollment) margin that matters, the endogenous (college choice) margin may be more important. This is especially true when subsidized public options are available at the lower end of the spectrum.

Inter-vivos transfers are an important source of college financing. Gale and Scholz (1994) discuss how these transfers are distributed empirically. The model distribution of these transfers, displayed in Figure 1.9 (where transfers are normalized by average labor earnings in a model period), resembles the data: it is heavily right-skewed, with a mass point at zero, and transfers are generally larger for those that go to college than for those that do not. Substantial grants also go to those who do not go to college: this is due to the model structure, in which there is only one moment for parents to act on their altruistic preferences.

Figure 1.9: Distributions of Inter-Vivos Transfers



Grants and subsidies have been discussed above. Student loans are the remaining form of student aid. The model's parameters that regulate Stafford loans were set to match overall subsidized and unsubsidized loan amounts, as well as their relative limits. The model also makes predictions on the fraction of students that take up loans. The fraction of students with subsidized Stafford loans was 37.3% in 2000, and the same fraction for unsubsidized Stafford loans was 21.2% (Abbott et al., 2013). These moments and their model equivalent are displayed in Table 1.7. Clearly, the model does not do a good job at matching loan uptake at the extensive margin. As a result, the model has too many students taking up loans, and too little loans per student, compared to the data. This would be a result of missing heterogeneity of eligibility, for example because eligibility in reality is tied to additional conditions, such as actual college expenses. Given the complexity of modeling such conditions, I have focused on matching overall loan availability at the expense of generating a realistic cross-sectional pattern of loan uptake. A similar comment applies to the modeling of inter-vivos transfers: the model presumably misses some sources of heterogeneity here too, for example in preferences, that would make financing needs more heterogeneous.

The overall role of government in the model is in line with the data. First, public spending on education, which in our model includes institutional grants, is about 1% to 1.5% of GDP depending on sources.²¹ Second, government expenditures amounted to 36.6% of GDP in

²¹For example, NIPA reports public expenditures of 0.91% of GDP for 2003, while institutional grants amounted to .16% of GDP (making for a total of 1.07%).

Table 1.7: Remaining Moments

	Model	Data
Share of students with Subsidized Stafford Loans	75%	37%
Share of students with Unsubsidized Stafford Loans	98%	21%
Public spending on education as % of GDP	0.77%	1.07%
Government expenditures as % of GDP	20.8%	37%
Weekly hours worked in college	33	>12
Weekly hours studied in college	15	<25
Frisch elasticity	1.20	>0.75
Wage premium	1.61	1.61

2003 according to the OECD. The model economy includes numerous sources of taxation but not all, so that it underestimates the size of the government somewhat. Model counterparts to both figures are reported in Table 1.7.

The final source of college financing is time use: students can choose how much time to spend working instead of studying or enjoying leisure. The ratio of time spent on education versus at work has already been targeted. As mentioned, quality data points on time use by students are hard to come by. Comparing the model to the data on this issue is difficult for a further reason: in the model, students are identified by enrollment, and therefore in principle include drop-outs, part-time students, students in two-year programs, and so forth, for an entire four year period. In spite of these issues, I produce model predictions of time use levels in Table 1.7. They are in line with sources other than those already reported from: Data from the National Center for Education Statistics (2017) lead to an estimate of 12.24 hours worked per week for a full time student (a lower bound for the model). According to the calculations by the Bureau of Labor Statistics based on data from the American Time Use Survey²², an average full time student spends 3.5 hours a day (or 24.5 hours a week) on educational activities (an upper bound for the model).

1.5.5 Labor Earnings

Labor earnings are front and center in the model presented here. How do these look in the model versus in the data? Some of the model's parameters have been tied down targeting age-earnings profiles from Huggett, Ventura, and Yaron (2011), as discussed above. Figure 1.10 shows these profiles in full, together with their model equivalents. The model matches the earnings life-cycle overall, although with a shortage of curvature. This is also the case in the work by Huggett et al., who take a similar approach. The model also generates too

²²<https://www.bls.gov/tus/charts/students.htm>, data are from the period 2011-2015.

strong an increase in the variance of log labor earnings.

Figure 1.10: Age-Earnings Profiles

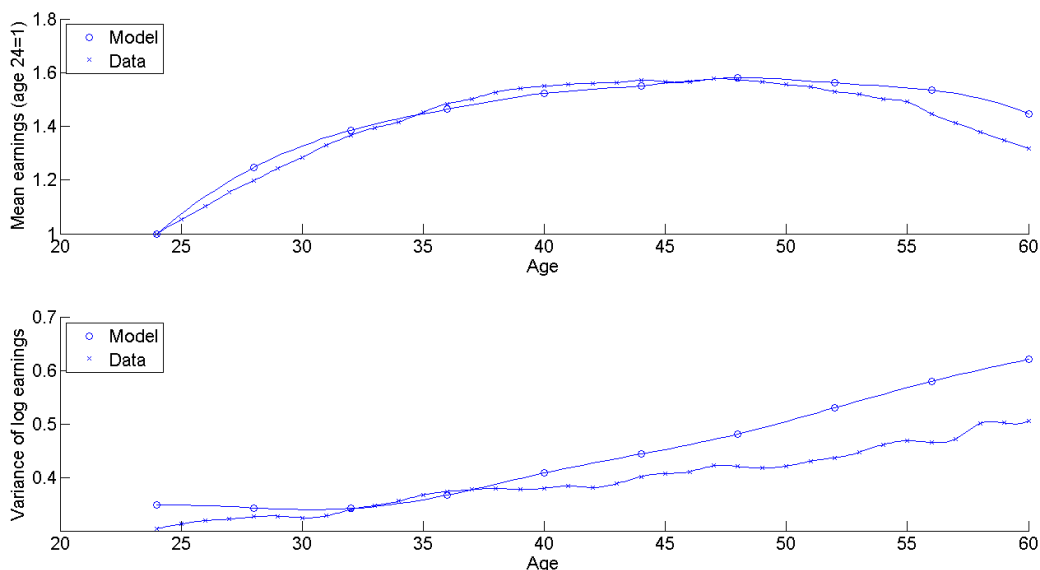
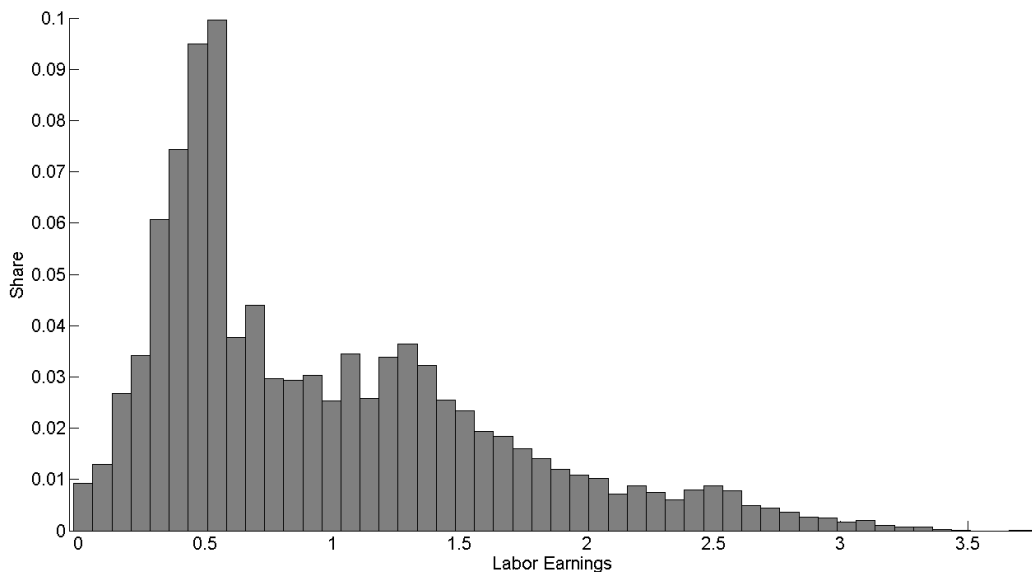


Figure 1.11 displays the model distribution of earnings, with average earnings normalized to one. The model distribution has much in common with the well-known empirical distribution, specifically that it is right-skewed and has a long right tail. Unreported results show a distribution of wealth with a significant population of indebted agents, and a large mass of agents with close to zero asset holdings. However, the wealth distribution does not produce the large right tail that is observed in data. This is due to the fact that the model does not include inheritances, nor inter-vivos transfers at ages other than the start of adult life. It also does not include a retirement period. Age-savings profiles also reflect the model structure: assets gradually deplete after an initial receipt of parental transfers, and the average agent starts building up savings again towards age 40. That build-up of assets is temporarily interrupted by inter-vivos transfers to children.

Frisch elasticities are typically used to measure the responsiveness of labor supply in macro models. Chetty, Guren, Manoli, and Weber (2011) argues for Frisch elasticities of 0.75 in macro models, but Keane and Rogerson (2012) argue that values well over one are more in line with the data once human capital is taken into account. The model's implied average Frisch elasticity is in that region, see Table 1.7.

Finally, the raw college premium at age 32, the ratio between the average wages of those with 16 years of education or more over the average wages of those with less, is 1.61. This is

Figure 1.11: Distribution of Labor Earnings



taken from a sample of the 2000 US Census, obtained from IPUMS.²³ The model generates a figure that is in line with this measure (see Table 1.7).

1.6 Counterfactuals

This section uses counterfactual policies to draw lessons from the model. It does so by comparing stationary equilibria of the model economy.

1.6.1 A Decomposition of Earnings Persistence

There are three intergenerational links in the model: ability is correlated over generations, parents can transfer money to their children, and education policies take into account parental wages. Education policies, in turn, consist of a subsidized public college, grants, and (Subsidized and Unsubsidized) Stafford student loans.²⁴

We can decompose the intergenerational persistence of earnings into its components. I start by removing each of the links. First, I remove all persistence from the ability transition matrix by assigning an equal probability to each destination level of ability (for each starting level).

²³The bottom 1% of incomes are dropped, as are those who work less than 40 weeks a year or less than 35 hours a week. The Census' own top-coding corrections are accepted as is.

²⁴Access to public college is not conditional on parental earnings, but I include this policy in my decomposition nevertheless.

Second, I set ω , the parameter that determines the extent to which parents internalize the well-being of their children, to zero. As a result, parents will no longer make any inter-vivos transfers. Third, subsidies to the public college, student grants, and loan limits are all set to zero. I then undo each of these changes, going in the same order. The results are displayed in Table 1.8. I display both the IGE and the rank correlation measure of intergenerational persistence, and also provide a break-down in percentages of the baseline model. Earnings are measured at age 36 for both generations, so that the measures are as clean as possible.

Table 1.8: Decomposition of Intergenerational Persistence

IGE		Rank Corr.		Ability Persistent	IV-transfers Full	Policies Full
0.0	(0%)	0.0	(0%)			
0.06	(38%)	0.08	(53%)	✓		
0.10	(63%)	0.10	(67%)	✓	✓	
0.16	(100%)	0.15	(100%)	✓	✓	✓

Roughly half of intergenerational persistence in the baseline model is driven by ability (as measured at age 18) alone. The remainder is split between inter-vivos transfers and education policies, with the former accounting for about a third and the latter for two-thirds. Here, it is important to note that when education policies are absent, inter-vivos transfers will adjust. That adjustment mechanism is important when measuring the effect of policies, as has previously been analyzed in work by Abbott, Gallipoli, Meghir, and Violante (2013).

Is the effect of transfers due to the level of the transfers or their distribution? These two component can be separated by an extra experiment in which all agents are given the average transfer (measured in the experiment with transfers but without policies). Unreported results show that this hardly changes persistence from the experiment without transfers. In short, it is not the level of the transfers that matters, but who gets them. Transfers matter because they help pay for college.

How important are different policies? We can investigate this by re-activating them one-by-one, starting with the public college, then grants, and then subsidized and unsubsidized loans. Table 1.9 presents the results. In each case, inter-vivos transfers and other choices have been allowed to adjust. It turns out that all policies increase persistence, with loans having the strongest impact. This is in line with the theoretical ambiguity discussed in Section 1.2, to which we will return below.

Again we can ask whether it is the level of these policies or their distribution that matters. Similar to the case for transfers, I run two extra experiments. The first gives the average grant (measured in the experiment with grants but without loans) to all that enroll in

college. The second gives the average loan (measured in the baseline model) to all that enroll in college. Giving average grants actually reduces persistence compared to the experiment without grants. The same applies to loans: giving average loans reduces persistence compared to the experiment without loans.

A look back at the distribution of aid in Figure 1.7 clarifies why. It is students from high-income families that receive most aid. This, in turn, is a consequence of their higher spending on education: in the data, grants are strongly related to college sticker prices. Redistributing aid evenly results in a policy that is more targeted at low-income students, which is where the positive effect on intergenerational mobility comes from. This brings us to the next lesson from these extra experiments: while policies modeled after actual policies increase persistence, there are some policies that reduce persistence. The effect of education policies on intergenerational mobility depends on the shape the policies takes. We will return to this below.

Table 1.9: Decomposition by Specific Policies

IGE		Rank Corr.		Public College	Grants	Loans Subs. Unsubs.	
0.10	(63%)	0.10	(67%)				
0.10	(63%)	0.11	(73%)	✓			
0.12	(75%)	0.13	(87%)	✓	✓		
0.15	(94%)	0.14	(93%)	✓	✓	✓	
0.16	(100%)	0.15	(100%)	✓	✓	✓	✓

Regarding the sources of persistence, it remains to note that the effect of education policies is significant in size: it is of the same order of magnitude as cross-country differences in persistence. It is also comparable to the effect of significant reductions of tax progressivity, which have been reported as important drivers of mobility elsewhere in the literature.

1.6.2 Removing Borrowing Constraints

Typically, the stated goal of government student loan schemes is to alleviate borrowing constraints that students would otherwise face.²⁵ I will now study how intergenerational

²⁵There are two reasons why the first welfare theorem breaks down in the environment this paper studies. First, government taxation drives a wedge between the private and social returns of labor, and therefore between the private and social returns to education. This creates complex optimal policies that have filled an extensive literature (see Section 1.1.2 for an introduction). I will not engage with the issue here. Second, markets are incomplete for two reasons: wages are subject to idiosyncratic shocks, and potential students have limited access to borrowing.

persistence would look if all borrowing constraints on students were alleviated. Specifically, I remove all education policies (including the public college), and then let all students borrow up to the natural borrowing limit in period zero. I treat this experiment as if the government is providing loans at market rates. Any net loss to the government from this change comes out of government expenditure.

Table 1.10: Removing the Borrowing Limit

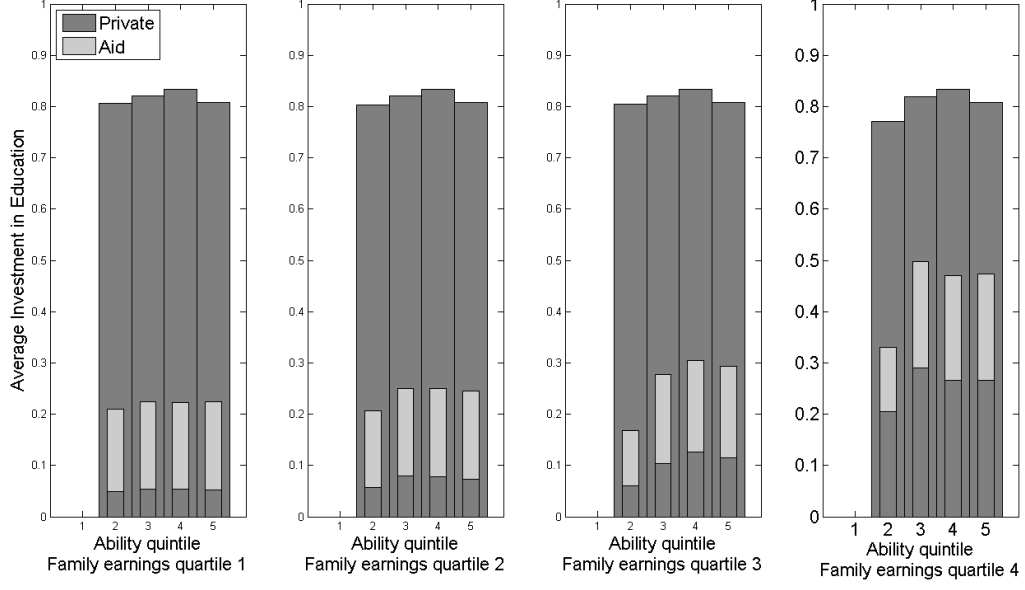
IGE		Rank Corr.		Baseline	Optimal
0.16	(100%)	0.15	(100%)	✓	
0.15	(94%)	0.14	(93%)	✓	✓

Table 1.10 reports the results of this experiment. Both measures of persistence move, but the effects are small. This is not because removing borrowing limits has no effects: education investment is much larger than in the baseline model. Instead, opposing effects on mobility, to be discussed below, cancel each other out.

What happens to educational choices when there are no constraints to borrowing? All potential students in the top four ability quintiles go to college, raising overall college enrollment to 80%. None of these students go to the public college, because their optimal investment level turns out to be much larger than the investment the public college makes. Investment quadruples for those that spent least before. Figure 1.12 shows investment by the same groups as before, and compares this to the baseline model. Differences in the investment level by ability almost disappear. The remaining college heterogeneity is much smaller, indicating that most college heterogeneity is due to financial constraints at the intensive margin.

Other model variables change as well. Average labor earnings increase by 14%. The college premium (measured at age 32) increases to 2.23 (a 39% increase). The composition of college financing changes: IV-transfers are now only used to redistribute assets to children and not to finance college. As a result they are now lower for those that go to college than for those that do not. Hours worked in college are almost zero. At the same time, hours studied (which are a complement to monetary investment) increase threefold. All of these suggest the presence of intensive margin constraints in college choice: in reality, students do work in college, and parental transfers do correlate with college spending. Finally, the patterns in Figure 1.4 (‘access’, ‘success’, and ‘mobility’) remain qualitatively similar, but access to high-investment colleges for students from low-income families does improve significantly.

Figure 1.12: Average Investment in Education: Unconstrained versus Baseline



The wider bars represent the unconstrained case.

The thinner bars refer to the baseline model, and display the same data as Figure 1.7.

1.6.3 Exploring the Effect of Education Policies

Let us return to the discussion in Section 1.2. There, equation 1.2 established the following decomposition of the IGE (of wages, ignoring labor supply):

$$\beta^{IGE} = \frac{Cor(\log h, \log h')Var(\log h) + \left[Cor(\log x, \log h') \frac{\sqrt{Var(\log h)}}{\sqrt{Var(\log x)}} \right] Var(\log x)}{Var(\log h) + Var(\log x)}$$

The point was the following: the impact of education policies is theoretically ambiguous, because they have two opposing effects. Relieving financial constraints makes children less dependent on their parents. This reduces both correlations in the above formula. However, the IGE is a weighted average of the two, so that changes in the weights are crucial. If policies increase the variance of log human capital, emphasis will shift to the first correlation, which is typically the larger of the two. As a result, measured persistence may go up rather than down.

Table 1.11 reproduces Table 1.1, but now provides model-based measurements of each of the terms. The constrained case refers to the model without any education policies. The unconstrained model is the case discussed above in which students can borrow at market rates up to the natural borrowing constraint.

The unconstrained policy reduces the persistence of human capital. This is as expected,

because removing borrowing constraints allows smart children from low-income families to invest in their education. The second term of the weighted average that makes up the IGE is small, and even zero in the case without constraints. Crucially, the variance of human capital more than doubles when the constraints are removed. This shifts the weight in the weighted average to the first term, which is larger than the second one. Human capital becomes a more important component of earnings persistence. As a result, removing the constraints has an adverse effect on intergenerational mobility.

Table 1.11: The effect of financial constraints on the IGE

	Unconstrained	Constrained
$Cor(\log h, \log h')$	0.46	0.48
$\left[Cor(\log x, \log h') \frac{\sqrt{Var(\log h)}}{\sqrt{Var(\log x)}} \right]$	0.00	0.03
$Var(\log x)$	0.08	0.08
$Var(\log h)$	0.16	0.07
β^{IGE} (wage rates)	0.29	0.24
β^{IGE}	0.15	0.10

It is perhaps here where the paper's main point is borne out best. We have seen above that education policies modeled after current US policy increase intergenerational persistence on the whole. Now we observe that there is a direct trade-off between intergenerational mobility and efficiency in a classical sense: when we consider the effect of a policy that removes borrowing constraints, earnings become more persistent across generations. Removing borrowing constraints achieves a Pareto improvement, but reduces intergenerational mobility.

A similar analysis can be used to compare the unconstrained policy (where loans are available at will, but no other education policies are in place) to the model's baseline. Table 1.12 shows that the unconstrained policy generates more mobility in human capital, and this time it does translate to more mobility in earnings. This highlights the extent to which grant schedules are currently skewed towards students from high-income families. Just making loans available does a better job at generating mobility (both in human capital and earnings) than do current policies.

Again the particulars of policies are important. Targeting grants and loans more towards low-income students would potentially lead to policies that increase intergenerational mobility. However, that does not take away from the trade-off this paper finds: policies that release constraints for high-income students are welfare improving too, but reduce intergenerational mobility.

The above analyses also highlight the importance of idiosyncratic shocks in the model. If

Table 1.12: The effect of financial constraints on the IGE

	Unconstrained	Baseline
$Cor(\log h, \log h')$	0.46	0.53
$\left[Cor(\log x, \log h') \frac{\sqrt{Var(\log h)}}{\sqrt{Var(\log x)}}\right]$	0.00	0.05
$Var(\log x)$	0.08	0.08
$Var(\log h)$	0.16	0.11
β^{IGE} (wage rates)	0.29	0.31
β^{IGE}	0.15	0.16

one were to directly measure the persistence of human capital, education policies would unambiguously reduce persistence. However, earnings are determined by more than human capital alone. Indeed, a long literature has established the importance of earnings shocks, both transitory and permanent, which are best thought of as all components that one might describe as ‘luck’. This paper’s parameterization of idiosyncratic shocks builds upon that literature.

What about the assumption that all individuals are subject to the same idiosyncratic shocks? Crucially, if education also increases the ‘luck’ component in earnings, then our model would overestimate the positive effect education policies have on persistence. A review of the literature on earnings shocks suggests that this is not the case: more educated individuals are subject to similar, or even somewhat smaller idiosyncratic income risks than are less educated individuals. For example, Meghir and Pistaferri (2004) report that the variance of unexplained earnings growth in their setup falls in education. Abbott, Gallipoli, Meghir, and Violante (2013) provide further evidence, and find little difference over education groups for persistent shocks, and a variance of non-persistent shocks that is smaller for the more educated. Finally, the findings of Blundell, Graber, and Mogstad (2015) for Norwegian administrative data do not contradict these conclusions. In short, the key modeling assumption on earnings shocks in this paper is conservative with respect to the main result.

1.7 Conclusion

Using a combination of theory and data, this paper has attempted to explore the connection between higher education policies and intergenerational mobility. It has shown that a human capital-based model does a surprisingly good job at explaining the persistence of earnings across generations, as well as its relation to college choice. We now know that the relation between higher education policies and intergenerational mobility is theoretically ambiguous. Going one step further, the parameterized model of this paper suggests that common higher education policies actually decrease intergenerational mobility.

This surprising finding is due to the fact that earnings do not just originate from human capital, but are at least also due to luck. Education policies that increase the mobility of human capital also decreases the importance of luck, thereby making earnings more persistent over generations overall. This has important implications for policy makers: First, they should not use higher education policies to target intergenerational mobility. Second, they should be careful not to interpret lower levels of intergenerational mobility as bad outcomes. As this paper has shown, they may just be the result of Pareto-improving policy. Second, to see whether education policies induce mobility in a desirable sense, one might want to measure the persistence of human capital, for example using test scores, rather than the persistence of earnings.

Several directions may be worth pursuing in further research. The paper introduces a model of colleges in a competitive landscape that other researchers may find useful. Some of the predictions that model makes have been discussed, but there are more. For example on the interplay between endowments and institutional grants. Connecting the model to data on colleges may prove fruitful.

The topic of college heterogeneity, which this work has shown to be important in a number of respects, may also warrant further exploration. In particular, government-based systems of higher education (as they are found in some European countries) often lack heterogeneity, which may have consequences for welfare, inequality, and intergenerational mobility.

As this paper has shown, college heterogeneity is also important for the study of financial constraints: even when all potential students can afford to go to *some* college (the *extensive* margin of college choice), they may not be able to enroll in the college that is best for them (the *intensive* margin of college choice). This finding would be complemented by direct empirical evidence. Indeed, most empirical literature on financial constraints looks at whether students are constrained from entering some college (see Lochner and Monge-Naranjo (2011a) for an overview), or are constrained when already in college (e.g. Stinebrickner and Stinebrickner, 2008). But to what extent are students constrained when choosing which college to attend?

1.8 Appendix

1.8.1 Computation

The computational procedure, for given parameter values, that produces the results in this paper proceeds as follows.

1. Guess the initial value function at independence V (because we interpolate between grid points, guesses at grid points are sufficient).
2. Solve the individual's problem. I elaborate on this below. This results in an updated function V .
3. Update V (i.e. repeat from step 1) until convergence.
4. Simulate households. This is done by randomly assigning initial states, and then simulating a household for many generations. Using a large number of households and a large number of generations per household, we arrive at a stationary distribution of the economy.

The individual problem is solved by backward induction. Because each life-cycle starts with a discrete choice, value functions will have kinks and (through inter-generational links that reach back many generations) so-called 'second-order kinks'. Thus, inter-temporal first-order conditions do not apply. Instead, the optimization to solve individual problems is done over time use using robust multi-level grid methods at each step of the life cycle. Leisure is assumed to always be interior, deviations from which are treated as a numerical imprecision.²⁶

The code for this procedure was written in Fortran90 and parallelized using OpenMPI.

²⁶The fraction of individuals that chooses zero work time is tiny, which is due to the absence of bequests in the model.

1.8.2 Student Aid in the United States

This appendix describes US tertiary education policies around 2003. Fuller (2014) provides a more detailed description of the history of US education policies.

Table 1.13: Education policies in 2003

Student Aid (as % of total, or in millions of 2012 USD)	
	2003
Grants (non-institutional)	
Pell	54%
Other Federal (mostly military)	19%
State	27%
<i>Total</i>	\$ 27,461
<i>% of GDP</i>	0.19%
 Institutional Grants	 \$ 22,470
<i>% of GDP</i>	0.16%
 Public Sector Loans	
Stafford, subsidized	44%
Stafford, unsubsidized	38%
PLUS	11%
Other Federal	4%
State and Institution Sponsored Loans	3%
<i>Total</i>	\$ 56,280
<i>% of GDP</i>	0.40%
 Private Sector Loans	 \$ 8,900
<i>% of GDP</i>	0.06%

Sources: author calculations; College Board (2013); CPI from the St. Louis FRED database; GDP from the World Bank's WDI.

Table 1.13 provides an overview of the student aid landscape in 2003. Government intervention generally consists of grants and loans. The largest uniform grant program is the Pell grant program, which provides grants to college students depending on financial need. Other programs are sizable but spread thin, with most of the money coming from states or military (including veteran) related programs. Institutional grants are of a similar order of magnitude as non-institutional grants. This uncovers a serious issue with using headline costs of college to calibrate models with an extensive margin: institutional grants are essentially discounts

to attending a college, and given their size the headline costs can hardly be taken to be the price of college. In addition, colleges often discount prices based on both financial need and merit. To account well for that complicated landscape, this paper relies on linked micro survey data on students and the colleges they go to.

Public sector loans, the other major policy instrument, largely consist of Stafford loans. These loans, in their subsidized version, provide student loans to students from lower income families at below market rates. Interest accrued during college is forgiven. Unsubsidized loans have higher interest rates and no accrual forgiveness, but are easily available to students regardless of family income. Subsidized and unsubsidized loans are subject to a joint limit, in addition to a separate limit on subsidized loans. Stafford loans are explicitly modeled and calibrated in this paper.

The other major loan programs are PLUS loans and Perkins loans. The availability of PLUS loans in practice strongly depends on parental credit scores, and are essentially a way for parents to transfer privately borrowed funds to children. This mechanism is separately present in the model through inter-vivos transfers, so that PLUS loans are not modeled explicitly. The Perkins loan program is small in size, and also not modeled.

Private student loan markets were small in 2003. Why this is so, not only in the United States but globally, is a topic of research in its own right. Here I put forward the following narrative: in the face of regular consumer bankruptcy regulation, private student loan markets are unlikely to develop at all (cf. Lochner and Monge-Naranjo (2011b)). Public student loans, presumably for the same reason, have historically been exempted from discharge in regular bankruptcy proceedings. Importantly, this exemption was extended to any non-profit entity in 1985, allowing many financial institutions to structure their loans such that they were immune to discharge (Consumer Financial Protection Bureau, 2012) and the private student loan market to develop.

Despite the discharge exemption, private student loans are not widely available: for example, 90% of undergraduate and 75% of graduate private student loans in 2012 were co-signed (MeasureOne, 2013). Without a co-signer, students typically lack a credit history that would allow them to take out a loan at competitive interest rates, but those that do take out these loans tend to get them at rates that are attractive compared to unsubsidized Stafford loans (Institute for Higher Education Policy, 2003). Because of their limited size, as well as their limited relevance to those students that are likely to face financial constraints in the absence of any loans, this paper does not model private loans.

Default on student loans are not part of this paper. It is precisely the exemption from discharge that makes this a less relevant issue for the purposes of this paper: students may

default, but then still have to repay their student loans. In fact, the College Board (2013) documents that over 90% of federal student loan dollars that enter default are eventually recovered.

Chapter 2

Polarization: A Supply-Side Mechanism

2.1 Introduction

2.1.1 Motivation

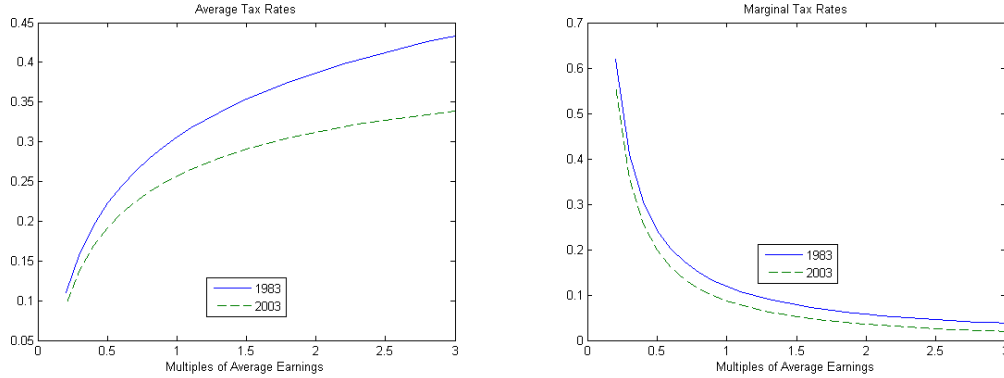
The effect of tax incentives on the formation of human capital have been studied extensively, both in positive (e.g. Guvenen, Kuruscu, and Ozkan, 2014) and normative (e.g. Stantcheva, 2017) contexts. This paper studies the incentive effects of taxation when human capital is multi-dimensional and the labor market is cleared in general equilibrium. Our results shed new light on relative movements of the earnings distribution. First, the presence of multi-dimensional skills can rationalize the relative unresponsiveness of low earnings to tax incentives. Second, we present a novel labor supply-driven mechanism for polarization of the earnings distribution to arise. The interaction between accumulation incentives and general equilibrium effects turns out to be key for the generation of non-monotone changes to the earnings distribution.

Tax progressivity is substantially different both across countries and over time. Since the 1970s, tax levels and progressivity in the United States have fallen dramatically. Guvenen, Kuruscu, and Ozkan (2014) estimate tax schedules on OECD data for the years 1973 and 2003. Both average ($t(y)$) and marginal ($t'(y)$) tax rates for the US are depicted in Figure 2.1.¹ Guvenen, Kuruscu, and Ozkan (2014) argue that high levels of taxes, and especially tax progressivity, play an important role in shaping the earnings distribution by reducing

¹Cross-country differences in tax progressivity and the relationship to inequality measures are plotted further below in Figure 2.6.

optimal human capital investments, particularly for the highly skilled. Taxes then compress the distribution of earnings. Guvenen, Kuruscu, and Ozkan argue that their mechanism explains changes in earnings inequality both across countries and over time, showing that it explains up to two-thirds of the change in the US wage premium between 1973 and 2003.²

Figure 2.1: Tax rates in the United States



While Guvenen, Kuruscu, and Ozkan (2014)’s results are suggestive, Skill-Biased Technological Change (SBTC) has been the main explanatory model of why inequality grew so strongly over recent decades in the United States (Katz and Murphy, 1992) and elsewhere (Berman, Bound, and Machin, 1998). According to the SBTC theory, increased supplies of highly educated labor keep the wage premium to education down, while technological change that is biased towards skilled labor continuously drives it up. In the 1980s the growth of educational attainment slowed down, explaining why inequality took off. In this type of theory, human capital is essentially considered two-dimensional, and general equilibrium effects are the main driver.

The perhaps primary challenge to both types of models is the ‘polarization’ phenomenon: the observation that starting in the 1980s jobs in the middle of the earnings distribution have seen less growth in wages and employment than those at the top or bottom, both in the US (Autor and Dorn, 2013) and across advanced economies (Goos, Manning, and Salomons, 2014). This coincided with growth in overall earnings inequality, i.e. a growing difference between the top and the bottom. Figure 2.2 displays these phenomena for the United States.

²All this is much in line with ideas from a public economics literature that considers how taxes are set optimally when human capital is endogenous (see Bovenberg and Jacobs (2005) for a static setting, and Stantcheva (2017) for a dynamic extension). In this line of research, human capital is considered one-dimensional, and general equilibrium effects on wages play no role. Recently, a literature has developed that considers the original Mirrlees problem when many types interact in general equilibrium. A recent contribution is by Sachs, Tsyvinski, and Werquin (2016). The formation of types (or human capital) has so far been taken as exogenous. The same applies to previous work by Teulings (2005), which provided a framework for tracing out general equilibrium effects across many types.

A number of explanations have been put forward to explain what is different about jobs in the middle of the income distribution, such as offshorability or competition from China or declines in unionization rates in the manufacturing sector, but consensus has formed around the view that these jobs have a higher degree of 'routineness', and are therefore more susceptible to automation (by machines, robots, and computers). In short, polarization of the labor market is seen as demand-driven, and attributed to exogenous technological forces. See Autor et al. (2010) for a review of this literature.

Figure 2.2: Wage Inequality Growth and Polarization in the United States



This paper takes an entirely different and complementary approach. It extends the analysis of incentive changes as in Guvenen, Kuruscu, and Ozkan (2014) to a multi-dimensional setting, in which there is also a role for the general equilibrium effects of Katz and Murphy (1992). Thus, it combines both of the established approaches to earnings inequality discussed above.³

We begin our analysis with a set of empirical results. Analyzing data on occupational skills from the Dictionary of Occupational Titles (DOT) combined with Census data, this paper establishes that there seem to be *two* relevant dimensions of job skills: cognitive and manual skills, both of which can be thought of as continuous variables. We find the importance

³A similar setup has been used by Guvenen and Kuruscu (2010) and Guvenen and Kuruscu (2012) to study skill-biased technological change. However, in both papers the general equilibrium wage effect is deliberately shut off by choosing a linear production technology. This precludes interaction effects between relative skill quantities and prices of the type we study in this paper.

of these skills to be heterogeneous over the distribution of earnings: manual skills play a relatively larger role at the bottom of the distribution, cognitive skills play a larger role at the top. Similar to Katz and Murphy, the cognitive skills coincide heavily with schooling decisions. We discuss the details of this analysis in Section 2.2.

Taken together, these points have important implications for the impact of tax incentives: First, those incentives are more relevant for those at the top of the income distribution than those at the bottom. This is because cognitive skills are subject to individual investment – and therefore incentives – to a much larger extent than manual skills, and because cognitive skills dominate at the top of the income distribution. Second, in general equilibrium, a change in the relative amount of cognitive skills may affect the relative prices of the two skill types and therefore individual earnings – a channel that is absent in models of one-dimensional human capital, but common in the literature on SBTC.

Motivated by these empirical findings, Section 2.3 sets out a simple life-cycle model in which earnings are derived from cognitive and manual skills, cognitive skills are subject to endogenous investment decisions and relative wages are determined in general equilibrium. Importantly, in our model of skills it is not different education levels that map into different skill types as is standard in the literature.⁴ Instead, both skill types are continuous, and one of them is formed by education. In Section 2.4 we use a simpler and more tractable representation of the model to study the effects of tax progressivity in our setting theoretically. We emphasize two implications: inequality and polarization.

First, our mechanism impacts *inequality*. Just like in the uniform human capital model, a decline in tax progressivity impacts the top of the income distribution more than the bottom, thereby increasing income inequality in absolute terms. This is because, in typical human capital models, the more able spend more time learning, and tax progressivity reduces the private gains from having learned earlier in life.

Two things are different in the multi-dimensional case, causing *polarization* to arise. First, the heterogeneous impact of taxes becomes much stronger, so that income inequality also increases in relative terms. This is because lower tax progressivity increases the relative supply of cognitive skills more than that of manual skills, thereby increasing the latter’s relative price. This lowers incentives to acquire cognitive skills. Second, the increasing relative price of manual skills, which are more important at the bottom of the income distribution, makes earnings at the bottom of the distribution even less sensitive to progressivity. If this effect is strong enough, it can even increase the wages of those at the bottom relatively more

⁴One notable exception is Lindenlaub (2017). She analyzes a matching model with multi-dimensional skills and shows how different rates of technological change between different skills can lead to polarization. Unlike in our model, she takes all skills as exogenously given.

than of those at middling levels of the distribution. In short, the tails of the distribution potentially respond relatively stronger to changes in tax progressivity, with a reduction in progressivity causing polarization in earnings. We discuss the underlying mechanism in further detail in Section 2.4.

The earnings distribution is likely subject to a multitude of economic forces and no single mechanism will be able to fully account for the changes that took place during the period in which polarization arose. We argue that our mechanism potentially contributes in addition to the existing mechanisms put forward in the literature, but do not claim it as the sole driver of the observed changes. In this paper, we attempt to study our supply-side mechanism in isolation. That limits the extent to which we can observe its empirical implications in data. Nevertheless, we include a qualitative comparison of the model’s macro-economic implications to data, both over countries and across time, in Section 2.5. The model predicts non-linear patterns of (relative) change in income distributions due to tax changes. While, as we discuss in detail, some of these patterns could also be generated by alternative setups with one-dimensional human capital, our theoretical exposition demonstrates that multi-dimensional models have a much easier time to account for such movements. Using OECD data on income distributions across countries and over time, we construct measures of income inequality and measures of tax progressivity across countries. Patterns that are easily accounted for by our multi-dimensional model are pervasive. We also discuss implications for changes over time, and the limitations present in verifying these.

Section 2.6 takes an enriched version of the model developed in Section 2.3 that can be solved numerically. Parameters are now tied down so that the baseline version of the enriched model matches relevant moments of the US economy. The model is then used to study the quantitative impact of a typical decline in tax progressivity.⁵ In order to do so, we use the decline of US tax progressivity since the 1980s that we discussed above as our experiment. We calibrate the model economy to match moments from the US economy in the early 2000s. We then compare the steady-state earnings distribution of this economy with the counter-factual tax progressivity of 1983 to the one in 2003 and calculate the rate of change in earnings.

The main goal of this exercise is to gauge the general quantitative “bite” of the human capital investment channel on changes in the earnings distribution, rather than wanting to account for the empirical change in earnings growth over the same period. This would require at the

⁵In this paper, we focus on tax progressivity. As Guvenen, Kuruscu, and Ozkan demonstrate, allowing for flexible labor supply makes tax levels a disincentive in the accumulation of human capital as well. In the context of a cross-country comparison, Guvenen, Kuruscu, and Ozkan find that differences in tax progressivity are a more important determinant of differences in inequality than are differences average tax rates, explaining our focus on the former.

very least taking into account the transition period as well as cohort composition effects, both of which our model is silent on. More generally, we are looking at our mechanism in isolation, whereas in reality several channels are likely to have played a role in the rise of polarization. The results from the experiment indicate that the model captures growth in overall wage inequality reasonably well. Most of the change comes out of the upper half of the income distribution, in line with the empirical evidence. The results further indicate that the polarization effect exists, and is quantitatively sizable but smaller than what we observe empirically. In conclusion, our mechanism has impact under quantitatively relevant variations in policy. We also argue why our estimate of the mechanisms quantitative implications might be seen as a lower bound.

Our results contribute to two separate literatures. First, they provide a natural explanation for the lack of response in the lower half of the income distribution to changes in the human capital accumulation incentive structure. Existing papers focusing on uniform human capital, such as Guvenen et al. (2014), lack explanatory power in this region of the distribution. By adding the general equilibrium relative price effect, our model complements the direct incentive effect studied in their paper with the general equilibrium price effect. The latter works primarily in the lower half of the distribution and helps to limit the increase in total inequality.

Second, existing theories of polarization are primarily labor demand driven. Autor and Dorn (2013) introduce a third 'routine' skill category and explain polarization through increasing automation of 'routine'-intensive tasks, reducing the demand for jobs located in the middle of the skill distribution. By adding general equilibrium relative price effects to the traditional skilled-unskilled dichotomy of the endogenous human capital literature, we are able to generate qualitatively similar changes in the earnings distribution without resorting to a third type of skill. Given the complexity of the earnings distribution, there are likely many underlying factors at work. Consequently, we see our channel as complementary to the skill-demand based explanations put forward in the existing literature.

Throughout, we emphasize that taxation is just one particular type of disincentive to human capital formation. In principle, there are many other distortions driving a wedge between public and private returns to education that differ across parts of the population. Two of the major trends in the 2nd half of the 20th century have been declines in gender and race based discrimination, both in education and the labor market. Hsieh et al. (2016) attribute about 25% of total economic growth in the US between 1960 and 2010 to changes in discrimination against women and minorities. It also seems reasonable to think that these trends correspond to a reduction in wedge progressivity, since discrimination seems like a particularly salient issue towards the top of the distribution. Thus, for the remainder of this paper, one may want

to think of ‘wedges’ more generally whenever we discuss taxes. Our quantitative results regard taxation only, so that investigating the role of discrimination for polarization is our main suggestion for further research. Section 2.7 concludes and provides further such directions for future research.

2.2 Manual and Cognitive Skills

We use data on the skill content of a number of occupations provided by the Dictionary of Occupational Titles. We analyze the structure of these data using a statistical technique (Principal Component Analysis) that allow us to reduce the dimensionality of the data and subsequently interpret them. We find that skills are best summed up by two dimensions: cognitive and manual skills, both of which are important. In order to map the skill content into the wage distribution, we link the DOT data to the Census. This also enables us to investigate how skill measures relate to other observables, such as education. It turns out that the cognitive component of skills is strongly correlated to measures of education, while the manual component is not. Below, we discuss our data sources in more detail. Empirical methodology and results are presented thereafter.

The main drawback of our type of analysis is that we look for underlying skill categories in the data *per se*, i.e. not in relation to the wages or schooling decisions we expect them to explain. Our results in first instance only aim to have explanatory power with regards to the questions and answers observed in the DOT. Several arguments speak for our approach nevertheless. First, the clear advantage of this approach is that our measures are in some sense still direct measures of skills, even if they are compounded and rely on analysts. Any explanatory power they have in our further analysis is not due to how we have produced them. Second, and related, the questions included in the DOT were included for a reason: because they were believed to be relevant measures of occupational skill. Last but not least, there is a related literature in which data on skills are directly related to wages. The most important reference in this regard is Yamaguchi (2012), who estimates a structural model of wage development in relation to unobserved skills using the same data on occupational skills as we do. He also finds two underlying skill factors to be of major importance, which he refers to as cognitive and motor tasks.

2.2.1 Data

DOT

We use the ‘Current Population Survey (CPS), April 1971, Augmented With DOT Characteristics and Dictionary of Occupational Titles (DOT)’, obtained from the IPCSR. This

version of the CPS was augmented with data on occupation characteristics from the 4th edition of the Dictionary of Occupational Titles (DOT). The 4th edition of the DOT is unique, in the sense that it is the final edition of the so-called ‘Analyst Database’. Over centuries, starting in the mid-1930s, the United States Employment Service led an effort to systematically document the skills required to perform a range of occupations. This was done by sending trained occupational analysts to job sites, where they would complete standardized questionnaires on occupation content. While the database was revised since, the focus after the 4th edition of the DOT shifted to the generation of O*NET data, which are based on surveys of employees and employers, and therefore much less suitable for comparison across occupations. Following much of the literature, we therefore choose to use the 4th edition DOT. The main advantage of using the augmented CPS database is that it provides us with numbers of workers per occupation in the original DOT occupation classification.

The 4th edition DOT provides information on 46 variables of skills needed for and characteristics of 3886 DOT occupations (some examples: Marine Architect, Die-Designer Apprentice, Weather Observer, Hypnotherapist). In the nationally representative CPS database that we use, we also have a proportion of the working population for each of these occupations. The 46 variables consist of the analyst’s answers to a wide variety of questions per occupation:

1. To what extent does the job relate to data, people, things? (3 questions)
2. What educational development is required (reasoning, mathematical, language, vocational)? (4 questions)
3. To what extent are aptitudes like intelligence important, or finger dexterity? (11 questions)
4. What temperaments relate to some occupation? (10 questions)
5. What are the physical demands of the job? (6 questions)
6. What physical environment does the job take place in? (7 questions)
7. To what interests does the job relate? (5 questions)

The survey includes clear and detailed instructions on how to answer these questions, making the answers comparable across occupations. In addition, many questions include a grading scale that seemingly targets the possibility of cardinal comparison. Whether cardinal interpretation is appropriate depends on the context, but clearly information is lost when not using these scales in some cardinal fashion. For example, aptitude ratings ask analysts to decide which quintile of the population an occupation falls into. On some questions, however,

analysts were only asked to indicate whether they are relevant to a job or not. In each case, we convert the answers provided into numerical values.

Because the questions in the database vary in type and topic, and their number is large, researchers typically make ex-ante decisions on which variables to use. For example, while some questions clearly relate to skills, others clearly do not (such as interest and environment variables). We try to keep any pre-selection to a minimum, and include the three categories of questions on the DOT when we perform Principal Component Analysis. These categories, comprising 18 questions, have in common that they must all be answered on a numeric scale that suggests some form of cardinal interpretation. (This is generally not true for the other categories: they are of the ‘yes or no’ type.) They also all clearly relate to skills, rather than the environment or personal characteristics of the typical person performing the job. We provide more detail on the 18 questions with our empirical results.

Census

We obtain a crosswalk between DOT occupation codes and Census 1990 occupation codes from the Analyst Resource Center (amongst others associated with the US Department of Labor). Whenever several DOT codes map into one census code, we take the average of component scores as the component score for that Census occupation. This procedure results in 452 occupations.

We use US Census data from IPUMS for all non-skill data (wages, hours worked, employment shares over Census occupations, education, and so forth), where we take a random sample of 50 thousand observations for each of the census years we use. Non-farm hourly wage rates are constructed by combining wage income and non-farm business income, following the example of the Census itself, and correcting for the number of weeks worked and the number of hours worked in a typical week. We reflate all wages to 2012 levels using the ‘CPI total items for the United States’ from the Federal Reserve Bank of St. Louis. We remove all occupations which are not present throughout our sample, as well as all farm occupations. The final number of occupations for which we have data in our sample is 308.

Population percentiles are obtained as follows. For each occupation in our sample we obtain mean hourly log wages \bar{w}^{occ} and the share of the population employed in the respective occupation x^{occ} . We sort occupations by their mean log hourly wage in 1980. We construct percentile employment shares x^{perc} and average wages \bar{w}^{perc} by mapping the occupation population shares into population percentiles. In particular, we assign to each percentile, the share of each occupation falling into the respective percentile using the 1980 population share per occupation. Doing this, we obtain the following conversion matrix C of dimensions

$\#occ \times 100$, which by definition maps the vector of occupation population share vector x^{occ} into population percentile vector x^{perc} :

$$C'_{1980}x^{occ}_{1980} = x^{perc}_{1980} \equiv \mathbf{1}.$$

By construction, the population percentiles obtained in this way are 1 in 1980. Percentile mean wages in 1980 are obtained by multiplying occupational mean wages with the conversion matrix:

$$\bar{w}^{perc}_{1980} = C'_{1980}\bar{w}^{occ}_{1980}.$$

To obtain the change in employment shares between 1980 and 2010, we first calculate how occupational employment changed in terms of 1980 percentiles and then compute the rate of change. In particular, we take the conversion matrix from occupations into percentiles in 1980 and multiply it with the occupation employment share vector in 2010 as follows:

$$\Delta_{2010-1980}x^{perc} = C'_{1980}x^{occ}_{2010} - x^{occ}_{1980}.$$

This calculation converts 2010 occupational population shares into the 1980 percentile bins. If for an occupation the employment share increased (decreased) relative to 1980, this will result in the respective population percentile to increase (decrease) as well.⁶ For the calculation of changes of wages we similarly multiply the 1980 conversion matrix with 2010 occupational mean log hourly wages and obtain the growth rate by taking the difference of the 2010 and 1980 percentile wages.

2.2.2 Empirics

PCA

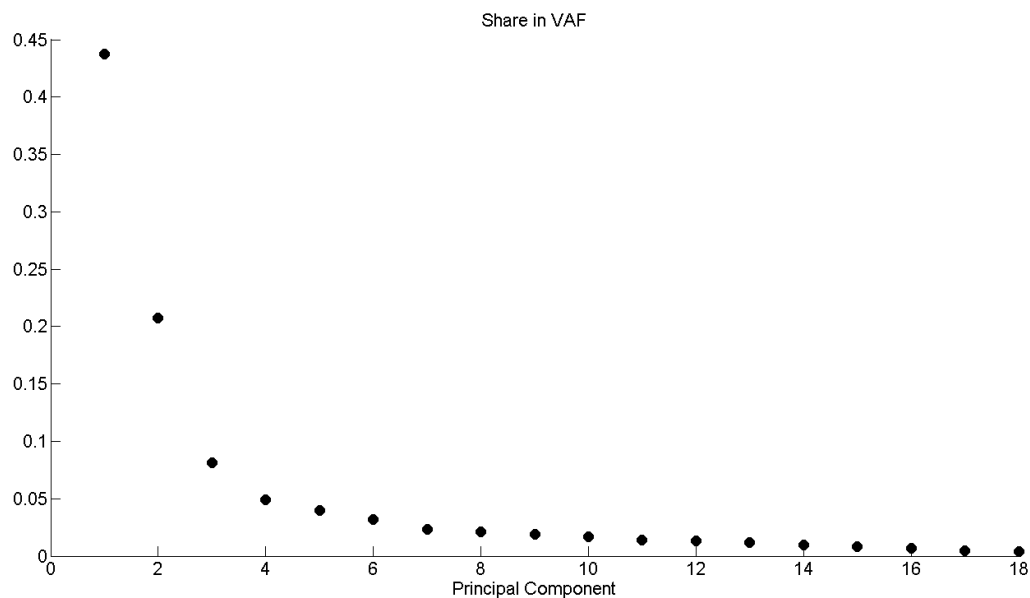
The leftmost column of Table 2.1 shows the labels of the 18 DOT questions we include in our analysis. This set is still large, so that we want to reduce it for more tractable empirical analysis. We think of these skills as *ex-ante* equally important indicators of underlying core skills, and want to find out what these underlying skills look like. One method that allows doing so is Principal Component Analysis (PCA).

PCA is a relatively standard technique for dimension reduction, that creates new variables by linear combinations of existing ones. Its objective is to maximize the variance of the new variable, which is called a principle component. Each subsequent component's vector of

⁶Note that by using this strategy, we are restricted to the analysis of occupations which are present in both 1980 and 2010. Thus, we remain silent on the effects of vanishing and newly appearing occupations on the aggregate wage and employment distribution. The procedure follows the approach taken by Autor and Dorn (2013).

weights to the variables is assumed orthogonal to the previous ones'. (To make this problem well-defined, variables are first standardized to have mean zero and standard deviation one, and the total weight given to each of them is restricted to be no more than one.) The optimality condition for this problem is a simple eigenvector-eigenvalue decomposition, which yields as many components (eigenvectors) as there are variables, with all components orthogonal to each other. The corresponding eigenvalues relate directly to the variance accounted for by each component. One can simply think of the components as new dimensions: the dimensions are rotated such that the first dimension explains as much variance as possible, thereafter the second, and so on. Thus, the components are identified up to sign and scaling. Variance accounted for by each component are displayed in Figure 2.3. Clearly, the first two components dominate the others in explanatory power: they jointly explain more than 60% of the variance in the data, while no other component explains more than 10%. The third component and further component do not seem to pick up a fundamentally different aspect of skills, but rather seem to modulate the first two. Full PCA results are included in Appendix 2.8.1.

Figure 2.3: PCA Scree Plot



What do the components look like? Table 2.1 shows the correlation between the first two components with the DOT skill measures over occupations (weighted by their share in employment). Those questions that are negatively correlated with the first component are highlighted. A brief study of the category groups with positive (negative) correlations with the first (second) component unambiguously leads to the conclusion that the first component relates to measures of cognitive ability, while the second component relates to physical skills.

Table 2.1: Component Correlations

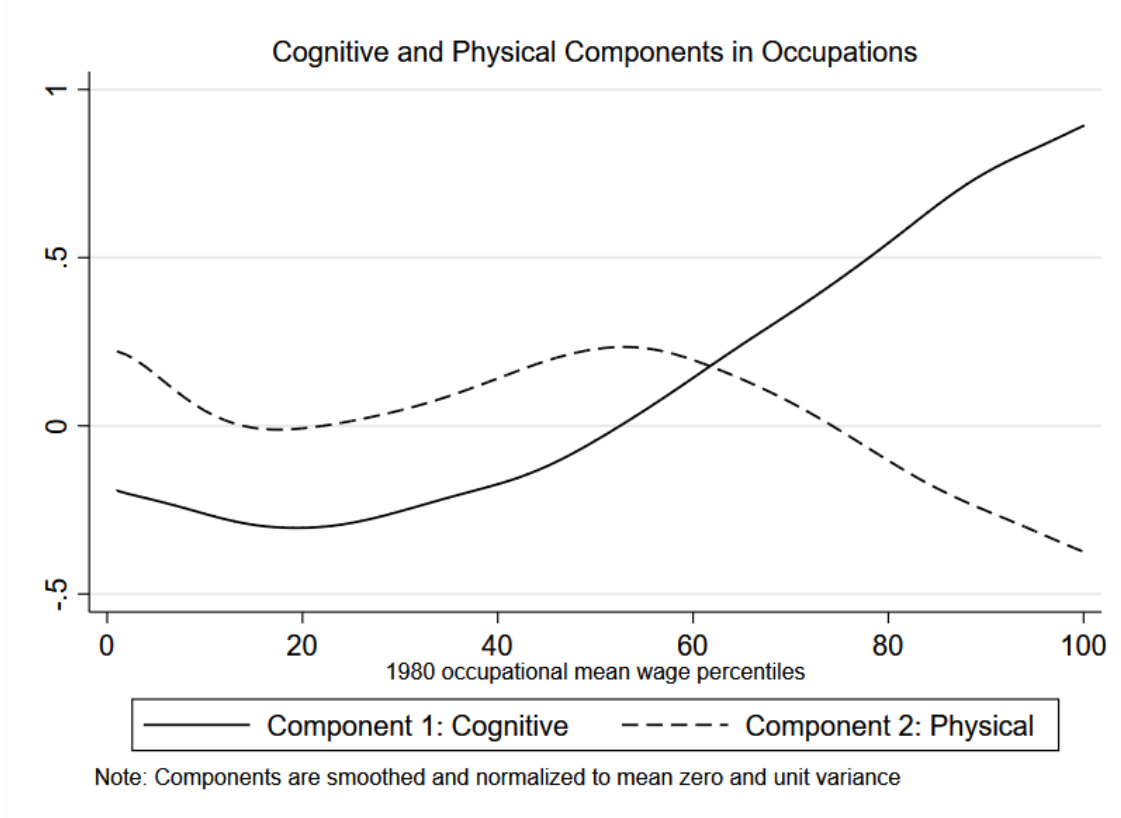
Variable	Component 1: ‘Cognitive’	Component 2: ‘Manual’
Data	0.49	-0.31
People	0.45	-0.27
Things	-0.34	0.12
GED Reasoning	0.47	-0.36
GED Mathematical	0.46	-0.39
GED Language	0.48	-0.38
Specific Vocational Prep.	0.36	-0.22
Intelligence	0.51	-0.44
Verbal	0.51	-0.49
Numerical	0.43	-0.62
Spatial	0.05	0.41
Form Perception	0.05	-0.09
Clerical Perception	0.53	-0.51
Motor Coordination	-0.44	-0.07
Finger Dexterity	-0.63	-0.14
Manual Dexterity	-0.57	0.21
Eye-Hand-Foot Coord.	-0.14	0.44
Color Discrimination	-0.28	0.46

The orthogonality assumption inherent to the method has the natural economic interpretation that these are truly different underlying skills: being good at one does not mechanically imply being good at another. At the same time, there can certainly still be correlation in abilities in the population of observed occupations. In fact, the correlation between observed occupation scores on the first two components is -0.25: those with more cognitive ability tend to be less able physically, and vice versa.

Covariates

We investigate how the results of the principle component analysis described in the previous relate to the wage distribution. Figure 2.4 plots the first two components over population skill percentiles. As one would expect, the cognitive component is of minor importance in the lower end of the skill distribution, and starts to increase in importance somewhere below the median. In contrast, the physical component is relatively flat for the lower half of the distribution. Above the median it continuously declines with increasing skill level.

Figure 2.4: Components over the Distribution



Taken together, the PCA results imply that the multi-dimensionality of human capital or skills captured in the DOT can be summarized in two main factors, which we call cognitive and manual. As expected, the physical skill is relatively more important in the lower half of the wage distribution, while the cognitive becomes increasingly important for the higher skilled occupations. We view this as an interesting result, as the skilled-versus-unskilled dichotomy has a long tradition in the analysis of human capital.

Empirically, the skilled-versus-unskilled distinction has often been proxied for by years of education or by comparing college educated workers to those without college education. Figure 2.5 compares the PCA cognitive component to these traditional skill measures. It plots the cognitive component alongside the average years of education and the share of college graduates across wage percentiles in the population. All three measures behave very similarly in a qualitative sense. They are systematically lower for the lower half of the distribution and exhibit a break around the median. In the right half, all three measures are quickly increasing. Table 2.2 presents correlations between these measures of education and the two components, confirming the results from Figure 2.5: Education strongly correlates with cognitive skills, but not with manual skills.

To sum up, the empirical results suggest that both manual and cognitive skills are impor-

Figure 2.5: Components and Education

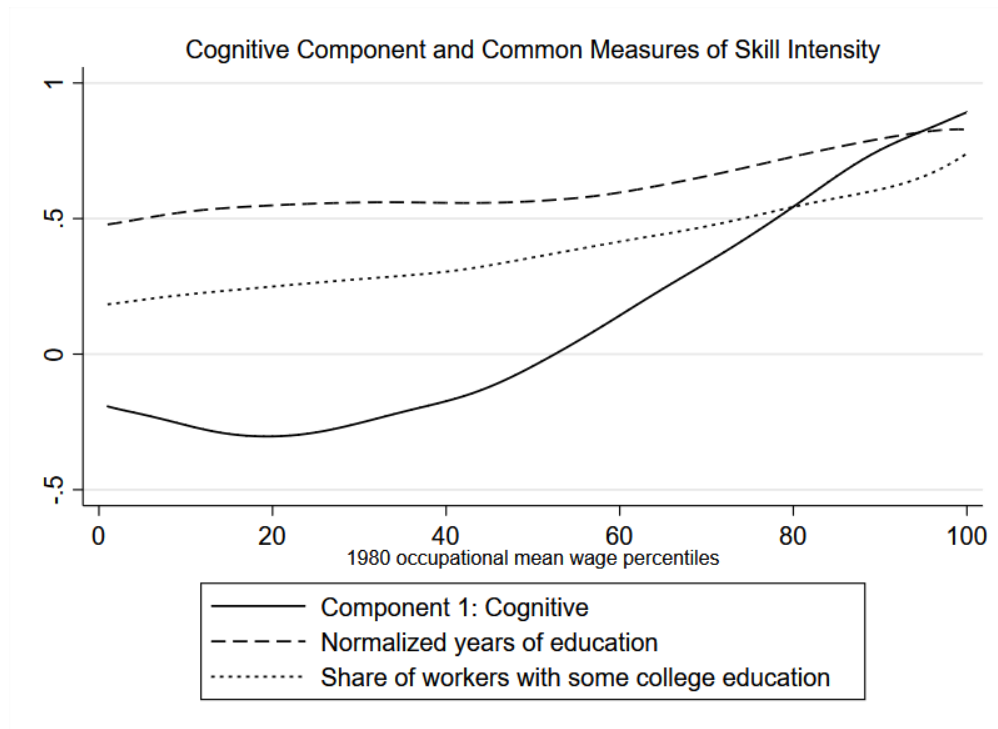


Table 2.2: Correlation between skill components and education measures

Correlations	‘Cognitive’	‘Manual’
Years of Education	0.97	-0.83
College Share	0.98	-0.70

tant, manual skills predominate in occupations at the bottom of the income distribution while cognitive skills predominate in the upper half. Furthermore, cognitive skills are highly correlated with traditional measures of schooling, which suggests that the cognitive skill is formed through education.

These observations have implications for the design of our theoretical framework. In particular, they suggest a model of skills where both manual and cognitive skills are important. This is much in line with the literature on Skill-Biased Technological Change, however there is one crucial difference: In the SBTC literature, it is commonly assumed that agents exclusively supply one type of labor, skilled or unskilled. Empirically, the most common strategy for mapping workers to skill types is by applying a cut-off for years of education above (below) which workers are categorized as skilled (unskilled). In contrast, the PCA results presented above suggest that both skills are continuous, i.e. workers in each occupation supply a *bundle* of cognitive and manual skills, rather than just one of the two. This precludes the mapping from years of schooling to skill categories. We will discuss the implications for our theoretical

framework in the next Section. Importantly, skill continuity allows for heterogeneity that is not given, but formed endogenously through schooling.

2.3 Model

In the remainder of the paper, we will theoretically and quantitatively explore how tax policy changes distort skill accumulation incentives and thereby impact the shape of the earnings distribution. Motivated by the empirical evidence presented in the previous Section, in this Section we set up a general environment in which earnings are derived from multi-dimensional skills. Individual skills are determined by innate ability and investment to different degrees. We describe the individual income maximization problem trading off time in school and time at work and discuss how progressive taxation distorts this trade-off.

2.3.1 Environment

Human capital accumulation A continuum of agents of mass one derives (pre-tax) earnings from two skills, manual m and cognitive s , quantities of which are measured by \mathcal{H}_m and \mathcal{H}_s , respectively. Human capital is accumulated by spending time in school at the beginning of an agent's life. Agents are active for one period of time, $t \in [0, 1]$ and begin their active period in school. They can leave at any time $x \in [0, 1]$ to begin working.⁷ Individuals are endowed with cognitive and manual abilities $\alpha = (\alpha_s, \alpha_m)$. Ability α is the only source of heterogeneity in the population. It is continuously distributed with pdf $f(\alpha)$ over a finite and positive support $[\underline{\alpha}, \bar{\alpha}]^2$. Based on the correlation between schooling measures and the cognitive component discussed in Section 2.2, we assume that the manual skill \mathcal{H}_m is innate, i.e. it cannot be accumulated and depends only on manual ability α_m . In contrast, the cognitive skill \mathcal{H}_s is subject to human capital accumulation or schooling. The efficiency of schooling time depends on individual's ability α , the amount of human capital already accumulated and the time spent in school, according to a schooling function $s(\mathcal{H}_s, \alpha, t)$. Cognitive human capital $\mathcal{H}_s(\alpha, x) > 0$, is assumed to be a continuous and twice differentiable function. Finally, we assume that $\frac{\partial \mathcal{H}_s(\alpha_s, x)}{\partial x} > 0$: accumulated human capital is a strictly positive function of time spent in school.

After quitting school, human capital stays constant for the remaining active time of the agents. Human capital accumulation thus follows a differential equation

$$\frac{\partial \mathcal{H}_s}{\partial t} = \begin{cases} s(\mathcal{H}_s, \alpha, t), & t \leq x \\ 0, & t > x \end{cases} \quad (2.1)$$

⁷When we discuss the effects of policy changes in this model, we will essentially be comparing steady states.

and the amount of cognitive human capital while working is given by the level of human capital at time x , $\mathcal{H}_s(\alpha, x)$.

Multi-dimensional earnings Individuals are assumed to supply both of their skills to the market simultaneously. Labor supply is assumed inelastic, and set to one. This implies that each individual supplies $\mathcal{H}_m(\alpha)$ units of manual labor and $\mathcal{H}_s(\alpha, x)$ units of cognitive labor when working. Instantaneous gross earnings y of an individual can now be described as

$$y = w_m \mathcal{H}_m(\alpha) + w_s \mathcal{H}_s(\alpha, x). \quad (2.2)$$

Here, w_m and w_s are wage rates for efficiency units of manual and cognitive skills, taken as given by the agents.

On the production side of the economy, final output is produced by an aggregate production function taking the total amounts of manual and cognitive skills in the economy as inputs,

$$Y = F(M, S).$$

Here, M and S are aggregate amounts of manual and cognitive skills in the economy and given by

$$S = \int_{\underline{\alpha}}^{\bar{\alpha}} \mathcal{H}_s(\alpha, x) f(\alpha) d\alpha$$

and

$$M = \int_{\underline{\alpha}}^{\bar{\alpha}} \mathcal{H}_m(\alpha) f(\alpha) d\alpha = 1,$$

in light of the above normalization of manual skills. We assume competitive input markets, thus wages w_m and w_s are given by their respective marginal products.

2.3.2 Individual Problem

Consumption prices are taken as the *numeraire*. Markets are complete, there are no sources of uncertainty, and a single asset completes the market: agents can save and borrow asset a without limit (except for repayment at $t = 1$) at the discount rate: $r = \tilde{r}$.⁸ Taxes are assumed to be paid instantaneously over the rate of income y . This is a crucial model ingredient: in reality, tax schedules are applied yearly, which is a short frequency compared to the length of the life cycle. In our continuous time model, we capture this by applying the tax schedule to the wage rate at any instance.

⁸We do not model the capital stock of the economy in general equilibrium, which yields the same results as an economy that is ‘small’ and open to capital only, or with an aggregate production technology that is linear in capital.

An individual's problem then looks as follows:

$$\begin{aligned}
& \max_{\substack{x \in [0,1], \\ \{c_t\}_{t \in [0,1]}}} \int_0^1 e^{-\tilde{r}t} \frac{c^{1-\sigma}}{1-\sigma} dt & (2.3) \\
& \text{subject to } \forall t : \\
& \frac{\partial a_t}{\partial t} = -c_t(1 + \tau_c) + a_t r \quad \text{if } t \leq x, \\
& \frac{\partial a_t}{\partial t} = y_t(1 - \tau_y(y_t)) - c_t(1 + \tau_c) + a_t r \quad \text{if } t > x, \\
& a_0 = 0, \quad a_1 \geq 0, \\
& c_t \geq 0.
\end{aligned}$$

Agents decide on the duration of their education and on the life-cycle profile of consumption and savings.

The government levies taxes on consumption and earnings, τ_c and $\tau_y(\cdot)$, in order to meet wasteful government spending target G . We assume that the earnings tax $\tau_y(\cdot)$ is governed by two parameters, responsible for average tax level ϕ and the degree of tax progressivity θ , $\tau_y(\cdot) = \tau_y(\cdot; \phi, \theta)$.

2.4 Tax Policy, Inequality Changes and Polarization

In this Section, we study the effect of changes in the tax policy, in particular changes to the tax progressivity θ on the shape of the earnings distribution. Investing in education enables individuals to achieve higher earnings in a shorter time span. Recall that in our framework, taxes are not applied to life-time income, but instead levied on instantaneous earnings, in order to resemble real-world income taxation. For a given life-time income, positive tax progressivity punishes higher per-period earnings relative to an earnings profile that spreads out lower earnings over a larger fraction of the life-cycle. Through this channel, tax progressivity directly influences the optimal choice of time in school, x^* and cognitive human capital, $\mathcal{H}_s(\alpha, x^*)$. Next, we will present a static version of the general framework above. We will then define the notions of inequality and polarization in this framework and provide conditions for the latter to arise in response to a policy change.

2.4.1 A simple static framework

Life-cycle problem (2.3) above has a straightforward solution. Since individuals can only decide between going to school or working full-time, the amount of human capital is fully determined by the time spent in school, x , and cognitive ability, α_s . Second, because markets are complete, agents smooth consumption and the choice of the optimal x is unconstrained.

As there are no other choices in the model, the agent now simply maximizes lifetime after-tax income with respect to time in school:

$$\begin{aligned} \max_{x \geq 0} \quad & (1-x)y(1-\tau(y; \phi, \theta)) \\ \text{subject to:} \quad & \\ & y = w_m + w_s \mathcal{H}_s(\alpha, x). \end{aligned} \tag{2.4}$$

For tractability, we assume that there is no heterogeneity in the manual skill, so that we can normalize its level to one, $\mathcal{H}_m = 1$. This implies that the aggregate amount of manual skill, M is also equal to one.

In the following, we will be interested in how changes in the tax progressivity θ shape the income distribution in this environment. As discussed, the optimal schooling decision will be directly governed by the degree of tax progressivity. From now on, we will therefore directly work with $h(\alpha, \theta) \equiv \mathcal{H}_s(\alpha, x^*)$ instead of human capital $\mathcal{H}_s(\alpha, x)$, where x^* is the argmax of individual income maximization problem (2.4).

2.4.2 Policy changes and earnings distribution

Our basic interest is the study of the effect of taxation on the earnings distribution in the presence of multi-dimensional skills. To this end, we will formalize the notion of earnings inequality and earnings polarization in our framework and provide conditions for either of them to arise. As will become clear below, earnings polarization is a special case of earnings inequality growth, with additional restrictions on the relative movements of earnings within the lower tail of the distribution.

As outlined above, we will link the changes in the earnings distribution reported by Autor and Dorn (2013) to changes in the tax incentives to relative skill supply studied by Guvenen et al. (2014). Thus, we will be interested in the relative (percentage change) effects of policy on earnings, since this is the theoretical equivalent of Figure 2.2. In particular, we will then be interested in how these policy effects differ across different parts of the income distribution. Our general object of interest is therefore given by

$$\frac{\frac{\partial y}{\partial \theta}}{y}.$$

It turns out, this can be easily decomposed into separate parts as follows:

$$\begin{aligned}
\frac{\left(\frac{\partial y}{\partial \theta}\right)}{y} &= \frac{\left(\frac{\partial \left(\frac{y}{w_s}\right)}{\partial \theta}\right)}{\left(\frac{y}{w_s}\right)} + \frac{\left(\frac{\partial w_s}{\partial \theta}\right)}{w_s} \\
&= \frac{\left(\frac{\partial \left(\frac{w_m}{w_s}\right)}{\partial \theta}\right) + \frac{\partial h}{\partial \theta}}{\left(\frac{w_m}{w_s}\right) + h} + \frac{\left(\frac{\partial w_s}{\partial \theta}\right)}{w_s} \\
&\equiv \underbrace{\frac{w'_\theta}{(w+h)}}_{\text{price effect}} + \underbrace{\frac{h'_\theta}{(w+h)}}_{\text{quantity effect}} + \underbrace{\frac{w_{s\theta}'}{w_s}}_{\text{level effect}}
\end{aligned} \tag{2.5}$$

In the above, percentage changes in income have been separated into three terms. The first two terms describe the potential trade-off policy created in a multi-dimensional model: on the one hand, policies can increase (or decrease) incentives to acquire human capital, which we call a *quantity effect*, but when they do so for all individuals this increases (decreases) the overall supply of learnable skills in the economy, which can decrease (increase) their relative price - a *price effect*. Both these terms would then move in the same direction, but their relative importance and strength depends on an individual's schooling responsiveness to policy. This responsiveness will in principle depend on the level of the ability parameter α , generating potentially non-linear effects of policy changes on income changes. The last term above affects all individuals equally in percentage terms. It arises because wage effects are described in skill premium terms, but a policy reform can also impact the overall productivity level in an economy - hence the name *level effect*.

Polarization in this environment arises if relative income changes in response to a policy change are stronger in the tails of the income distribution than in the center of the distribution. Since in the model, income is entirely determined by ability, this is equivalent to comparing income responses for different ability levels. Formally, inequality growth and polarization in response to tax policy changes can be defined in terms of relative income changes as follows.

Definition 2. *Inequality Change and Polarization. Inequality change exists in response to a policy change in θ if for $\underline{\alpha}$ and $\bar{\alpha}$ the following inequality holds:*

$$\left. \frac{\left(\frac{\partial y}{\partial \theta}\right)}{y} \right|_{\alpha=\bar{\alpha}} < \left. \frac{\left(\frac{\partial y}{\partial \theta}\right)}{y} \right|_{\alpha=\underline{\alpha}} < 0. \tag{2.6}$$

Polarization exists if in addition to equation (2.6), for some $\hat{\alpha} \in (\underline{\alpha}, \bar{\alpha})$ the following holds as well:

$$\left. \frac{\left(\frac{\partial y}{\partial \theta}\right)}{y} \right|_{\alpha=\underline{\alpha}} < \left. \frac{\left(\frac{\partial y}{\partial \theta}\right)}{y} \right|_{\alpha=\hat{\alpha}} < 0. \tag{2.7}$$

Definition 2 restates income inequality growth and polarization in response to a decline in tax progressivity θ in concise terms. First, the inequality aspect, i.e. high income individuals pulling away even further from the rest of the population, requires a stronger relative income response to a policy change for high ability individuals than low ability individuals. Since this effect is negative for higher levels of tax progressivity θ , the response will be more negative for high ability individuals. Second, the non-monotonicity in the lower tail of the income distribution distinguishes polarization from general trends in overall inequality: low-income individuals are able to partially catch-up to medium income individuals, while overall inequality still increases.

2.4.3 Conditions for Inequality Growth and Polarization

Our reformulation of the life-cycle problem (2.3) as income maximization problem (2.4) allows us to establish conditions for polarization to arise in our framework. After presenting those conditions, we will show their sufficiency for polarization to arise. The intuition for this result can also be seen from reformulating the decomposition (2.5) in terms of tax policy elasticities as follows:

$$\varepsilon_{\theta}^y = \varepsilon_{\theta}^w \frac{w}{w + h(\alpha)} + \varepsilon_{\theta}^h(\alpha) \frac{h(\alpha)}{w + h(\alpha)} + \varepsilon_{\theta}^{LE}. \quad (2.8)$$

Equation (2.8) rewrites the total relative earnings elasticity in terms of weighted elasticities of the price and the quantity effect, both of which are functions of α and the level effect, which is independent of α . Polarization can arise because the elasticity of the price effect has the same size for all abilities, while the elasticity of the quantity effect potentially grows in α . In addition, the weights also change in ability, since for higher ability the share of income generated from cognitive human capital increases. Depending on the shape of the change of elasticity of the quantity effect, we can have non-monotone changes in relative income across abilities. In particular, if $\varepsilon_{\theta}^h(\alpha)$ is small in absolute magnitude for small and medium α , the growing weight on the second, quantity elasticity term may initially decrease the absolute magnitude of overall elasticity as we move along the earnings distribution. Only once $\varepsilon_{\theta}^h(\alpha)$ is large enough in absolute magnitude will the absolute magnitude of the overall elasticity begin to grow in earnings.

In the following, we will make this reasoning more precise by first laying out the assumptions sufficient for polarization to arise, and then go through the precise mechanism. The main purpose of this exercise is to clarify and formalize the intuition just laid out.

Definition 3. Define $\hat{\alpha} \in (\underline{\alpha}, \bar{\alpha})$ as any α that generates a local extremum in the tax elasticity of income,

$$\left. \frac{\partial \varepsilon_{\theta}^y(\alpha)}{\partial \alpha} \right|_{\alpha=\hat{\alpha}} = 0.$$

Inequality change and polarization in Definition 2 were defined for an arbitrary interior $\hat{\alpha}$. Definition 3 restricts, as we will see below, the interior $\hat{\alpha}$ to the ability level that under the below assumptions minimizes (in absolute terms) the income elasticity $\varepsilon_{\theta}^y(\alpha)$.

Assumption 1. *Shape of the human capital elasticity. Human capital elasticity $\varepsilon_{\theta}^h(\alpha)$ behaves relative to the relative wage elasticity ε_{θ}^w as follows for different ability levels:*

- *The human capital quantity elasticity is increasing and convex in ability level α in absolute terms, $\frac{\partial \varepsilon_{\theta}^h(\alpha)}{\partial \alpha} < 0$ and $\frac{\partial^2 \varepsilon_{\theta}^h(\alpha)}{\partial \alpha^2} < 0$, and at the lower bound approximately zero: $\frac{\partial \varepsilon_{\theta}^h(\underline{\alpha})}{\partial \alpha} \approx 0$.*
- *For abilities $\alpha \leq \hat{\alpha}$, the human capital elasticity is lower in absolute terms than the elasticity of relative prices, $\varepsilon_{\theta}^h(\alpha) - \varepsilon_{\theta}^w > 0$.*
- *For high ability individuals, the human capital elasticity is higher in absolute terms than the elasticity of relative prices, $\varepsilon_{\theta}^h(\bar{\alpha}) - \varepsilon_{\theta}^w < 0$.*

Assumption 1 states that for low ability individuals, the relative price elasticity is stronger than the quantity elasticity. The quantity elasticity is increasingly growing in ability and for abilities high enough, it becomes larger than the relative price elasticity in absolute terms.

Assumption 2. *Human capital function. Human capital accumulation is strictly convex and positive in α , $h'_{\alpha}, h''_{\alpha} > 0$. In addition, for all $\alpha \in [\underline{\alpha}, \bar{\alpha}]$, the following restriction on the shape of the optimal human capital quantities holds:*

$$\frac{1}{w + h} < \frac{2(h'_{\alpha})^2}{h''_{\alpha\alpha}}.$$

Recall from above that $h(\alpha, \theta)$ is the cognitive human capital resulting from the optimal schooling decision of the agent, $h(\alpha, \theta) \equiv \mathcal{H}_s(\alpha, x^*)$. Therefore, Assumption 2 effectively imposes restrictions shape of the schooling technology. In particular, it is required that the human capital is convex in ability, but cannot increase too quickly – h''_{α} has to be sufficiently small.

Result 1. *Given Assumption 1, inequality changes as defined in Definition 2 occur in response to a change in tax policy θ . If in addition Assumption 2 also holds, polarization as defined in Definition 2 occurs as well.*

Proof: For the inequality change part, we will show that income elasticity $\varepsilon_{\theta}^y(\alpha)$ is larger in absolute terms for $\bar{\alpha}$ than for $\underline{\alpha}$. For the polarization part, we will show that under Assumption 2, there is a unique $\hat{\alpha}$ as defined in Definition 3 and this $\hat{\alpha}$ is the argmax of the maximum of $\varepsilon_{\theta}^y(\alpha)$ (minimum in absolute terms).

Inequality change: To show that $0 > \varepsilon_\theta^y(\underline{\alpha}) > \varepsilon_\theta^y(\bar{\alpha})$, we will first show that $\varepsilon_\theta^y(\underline{\alpha}) - \varepsilon_\theta^{LE} > \varepsilon_\theta^w$ and second that $\varepsilon_\theta^y(\bar{\alpha}) - \varepsilon_\theta^{LE} < \varepsilon_\theta^w$. To show the former, consider

$$\begin{aligned} & \varepsilon_\theta^y(\underline{\alpha}) - \varepsilon_\theta^{LE} > \varepsilon_\theta^w \\ \Leftrightarrow & \varepsilon_\theta^w \frac{w}{w+h(\alpha)} + \varepsilon_\theta^h(\alpha) \frac{h(\alpha)}{w+h(\alpha)} > \varepsilon_\theta^w \\ \Leftrightarrow & \varepsilon_\theta^w \frac{w}{w+h(\alpha)} + \varepsilon_\theta^h(\alpha) \frac{h(\alpha)}{w+h(\alpha)} > \varepsilon_\theta^w \frac{w}{w+h(\alpha)} + \varepsilon_\theta^w \frac{h(\alpha)}{w+h(\alpha)} \\ \Leftrightarrow & \varepsilon_\theta^h(\underline{\alpha}) > \varepsilon_\theta^w, \end{aligned}$$

where the last inequality holds by Assumption 1. For the high ability case $\varepsilon_\theta^y(\bar{\alpha}) - \varepsilon_\theta^{LE} < \varepsilon_\theta^w$, a similar argument holds. Together, this implies that $\varepsilon_\theta^y(\underline{\alpha}) - \varepsilon_\theta^{LE} > \varepsilon_\theta^w > \varepsilon_\theta^y(\bar{\alpha}) - \varepsilon_\theta^{LE}$. $\varepsilon_\theta^y(\underline{\alpha}) > \varepsilon_\theta^y(\bar{\alpha})$ is implied by the last inequalities, establishing inequality change as defined in equation (2.6) from Definition 2.

Polarization: To show that the unique global maximum of $\varepsilon_\theta^y(\alpha)$ is at $\hat{\alpha}$, we show first that $\hat{\alpha}$ is the only extremum, and second that the first derivative is strictly larger (smaller) than zero for all α smaller (larger) than $\hat{\alpha}$. The first derivative of $\varepsilon_\theta^y(\alpha)$ with respect to α is given by

$$\frac{\partial \varepsilon_\theta^y(\alpha)}{\partial \alpha} = (\varepsilon_\theta^h(\alpha) - \varepsilon_\theta^w) \frac{wh'_\alpha}{(w+h)^2} + \frac{\partial \varepsilon_\theta^h(\alpha)}{\partial \alpha} \frac{h}{w+h}. \quad (2.9)$$

Therefore, $\frac{\partial \varepsilon_\theta^y(\alpha)}{\partial \alpha} > (<) 0$ boils down to

$$(\varepsilon_\theta^h(\alpha) - \varepsilon_\theta^w) \frac{wh'_\alpha}{(w+h)^2} > (<) - \frac{\partial \varepsilon_\theta^h(\alpha)}{\partial \alpha} \frac{h}{w+h}.$$

For the left-hand side, Assumption 1 implies that $(\varepsilon_\theta^h(\alpha) - \varepsilon_\theta^w)$ is strictly declining in α , positive for $\underline{\alpha}$ and negative for $\bar{\alpha}$. Define $\tilde{\alpha}$ as the α such that $\varepsilon_\theta^h(\tilde{\alpha}) - \varepsilon_\theta^w = 0$. Note that by Assumption 1 $\bar{\alpha} > \tilde{\alpha} > \hat{\alpha}$ holds. Assumption 2 implies that $\frac{wh'_\alpha}{(w+h)^2}$ is strictly decreasing in α and strictly positive. Together this implies that the left-hand side is strictly declining in α for $\alpha < \tilde{\alpha}$, positive for $\alpha < \tilde{\alpha}$ and negative for $\alpha > \tilde{\alpha}$. For the right-hand side, Assumption 1 implies that $-\frac{\partial \varepsilon_\theta^h(\alpha)}{\partial \alpha}$ is ≈ 0 for $\underline{\alpha}$ and strictly increasing and positive for all $\alpha > \underline{\alpha}$.

Taken together, this implies that $\frac{\partial \varepsilon_\theta^y(\underline{\alpha})}{\partial \alpha} > 0$ and $\frac{\partial \varepsilon_\theta^y(\bar{\alpha})}{\partial \alpha} < 0$. Furthermore, since the left-hand side is strictly declining while positive and the right-hand side strictly increasing and positive, there exists exactly one α for which $\frac{\partial \varepsilon_\theta^y(\alpha)}{\partial \alpha} = 0$. This proves the existence of a unique $\hat{\alpha}$ as defined in Definition 3. Since for all $\alpha < \hat{\alpha}$, we have that $\frac{\partial \varepsilon_\theta^y(\alpha)}{\partial \alpha} > 0$ and for all $\alpha > \hat{\alpha}$, we have that $\frac{\partial \varepsilon_\theta^y(\alpha)}{\partial \alpha} < 0$, $\hat{\alpha}$ is a global maximum. Since $\hat{\alpha} \in (\underline{\alpha}, \bar{\alpha})$, this establishes polarization as defined in equation (2.7). \square

The aim of this Section has been to detail conditions on the price and quantity effects for earnings polarization to arise in our framework with two types of skills and general

equilibrium price effects. We show that depending on the shape of the elasticities, polarization can arise in our framework of two-dimensional skills and general equilibrium skill price effects. In the following, we will first present some reduced-form cross-country evidence for our mechanism in the next Section, and then try to quantify the economic importance of this supply-side channel in a richer version of our model in Section 2.6.

2.5 Models versus Data

2.5.1 Across Countries

Tax systems differ in progressivity across countries (Figure 2.6). Our model makes clear predictions on the role of progressivity in income inequality: more progressive tax systems produce less inequality as measured by relative earnings in the income distribution. This is driven by changes in the upper half of the income distribution, while inequality in the bottom half varies little with tax progressivity. Neither prediction is made by a model with one-dimensional skills, nested in our setup.⁹ We now investigate these predictions in cross-country data, for which we need measures of tax progressivity and relative earnings inequality.

Coen-Pirani (2017) sets forth a method to obtain measures of tax progressivity from OECD data, which works as follows: If we assume that both gross and net earnings are log-normally distributed and that taxes follow the functional form assumed above, measured Gini coefficients of gross and net earnings can be used to back out an estimate of θ . A panel data set of Gini coefficients is available from the OECD Income Distribution Database. We use data on the working age population (ages 18 to 65), using the income definition that the OECD followed until 2011 for better availability and comparability of data. Data are available for about 30 OECD member countries, covering a period from the mid 1970s to 2015. Coverage is thin for earlier years, but improves towards the end of the sample. Because the panel is rather unbalanced, we average the resulting measures of tax progressivity for the years 2010–2015, and use this as a cross-section of country-level tax progressivity.

Also available from the OECD is an unbalanced panel of relative earnings inequality measures across countries and over time. The underlying population are full-time employees of either gender. These include the earnings ratio of the 90th percentile cut-off to the 50th percentile cut-off, or *90-50 ratio*, and the same for the 50th and 10th percentile, the *50-10 ratio*. While these two measures describe relative inequality above and below the median, the resulting *90-10 ratio* measures inequality. We choose to use these measures because their movement

⁹We think of tax systems as exogenous to the remainder of the economy. If tax progressivity is in some way a response to higher earnings inequality, this would counteract our mechanism and make it harder to find a correspondence between theory and data.

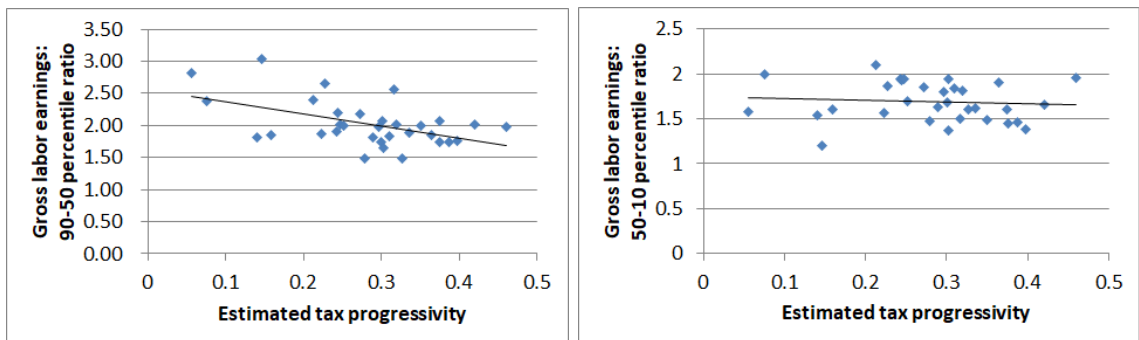
has a close correspondence to what we consider in our theoretical exposition: if relative (percentage) changes are the same across the distribution, then these measures will remain unchanged with tax progressivity. We again average over the years 2010–2015. The overlap between the two datasets consists of 32 countries.

Next, we estimate the (linear) impact of tax progressivity on earnings inequality at different points in the distribution. Results of OLS regressions of the latter on the former are displayed in Table 2.3. Figure 2.6 presents the results graphically. Tax progressivity is generally associated with a reduced relative earnings inequality. For the 90-10 ratio, the slope is statistically significant at the 5% level. For the 90-50 ratio the slope is statistically significant at the 1% level. For the 50-10 ratio, the slope is not statistically significant, even at the 10% level. While tax progressivity has quite some explanatory power in the upper half of the distribution, as measured by the R^2 , this is not true for the lower half of the distribution.

Table 2.3: Regression results

	<i>Inequality measure</i>		
	90-10 ratio	90-50 ratio	50-10 ratio
Progressivity ϕ_1	-1.17 (0.54)	-1.92 (0.59)	-0.19 (0.41)
Constant	Yes	Yes	Yes
Observations	32	32	32
R^2	0.14	0.26	0.01

Figure 2.6: Tax progressivity and Inequality across Countries



All these results align very well with our model prediction, even as we are looking at simple linear relations. In a more extensive quantitative exercise, Guvenen, Kuruscu, and Ozkan (2014) analyze the responses of a one-dimensional model of human capital to changes in tax progressivity. Their model has several added features, such as flexible labor supply, and a more flexible functional form for average tax rates. While they show that their model does well in accounting for 90-10 ratios, it is less succesful in disentangling 50-10 ratios. Our

analysis suggests that this is due to the multi-dimensional nature of skills, which is most relevant to the bottom half of the distribution. It also suggests that the productivity effect of taxes on human capital may be mitigated by general equilibrium effects.

2.5.2 Over Time

Tax progressivity in the United States has declined dramatically since the 1970s (Figure 2.1). That same observation applies to many other countries (Guvenen, Kuruscu, and Ozkan, 2014). What implications would this have had for other observables, in particular changes to the shape of the earnings distribution? The literature review above already shows that the earnings distribution is shaped by a number of different forces. That fact significantly limits the extent to which we can verify the direct impact of tax changes on inequality through our mechanism empirically. Nevertheless, we attempt to provide a qualitative discussion.

Educational decisions are decisions for the long run. Agents expectations of future policies are therefore key to the empirical mechanism we describe, and observed transitions may be slow. In any case, one would expect that younger cohorts react more strongly to incentives than older ones, so that empirically it should be the younger cohorts that cause polarization. This is indeed what the empirical literature finds. Cortes, Jaimovich, and Siu (2017) document that the fall in what they call ‘routine’ occupations in the middle of the distribution can be largely attributed to two groups: young and prime-aged men with low levels of education where it comes to ‘routine manual’ occupations, and young and prime-aged women with intermediate levels of education where it comes to ‘routine cognitive’ occupations. In terms of age structure, this lines up well with the implications of our mechanism. While our model does not speak to gender *per se*, the gender differences these authors highlight underline our main suggestion for further research: changes in labor market discrimination may be important. We will come back to this in more detail in the concluding Section 2.7.

Implications of changes in tax incentives for wages are summarized by rising inequality, in gross wages and even more so net of taxes, and polarization. While these phenomena can also be observed in the data, their underlying components cannot. This is because observed wages are the product of human capital quantities and prices. Did wage inequality grow due to greater differences in human capital or due to rising prices for the highly skilled? Our mechanism would suggest the former. The theory of Skill-Biased Technological Change on the other hand takes growth of educational attainment at face value as a measure of human capital quantities, and interprets its slow-down as a reason for rising prices for the highly skilled.

Separating human capital quantities from their price is a central empirical challenge in the labor literature and existing evidence is scarce. One approach is to identify an age in the

life-cycle at which human capital is unlikely to change much, and attribute wage changes at that age to changes in the price of human capital. This is the approach followed by Bowlus and Robinson (2012). These authors do not find large changes in prices at all, attributing changes in the wages of different educational groups to changes in human capital. This would be more in line with our mechanism than for example SBTC, although to cause polarization on its own our general equilibrium effect would require a growing relative price of manual versus cognitive skills. Price estimates by such skill types are unfortunately unavailable.

Similar caveats apply to direct measures of human capital (such as schooling attainment), measured skill premia, and before and after tax returns to schooling. While our model makes predictions for each of these, it is not clear what is the relevant empirical counterpart. A number of possible comparisons are further complicated by the fact that our model is not a growth model, so that it cannot account for longer-run trends in these data.

Finally we return to our initial comment: our mechanism is unlikely to have been the only relevant change during the period. Other explanations focus on secular technological developments that have shaped the wage distribution through changes in labor demand. These explanations are complementary to ours as long as relative prices of skills move in the same direction as in our model. That holds for the literature that describes how middling ‘routine’ occupations are more prone to automation. The same applies to papers that explain the growth of service occupations at the bottom of the distribution through changes in demand. SBTC fits our model less well, since it starts with the assumption that it is prices of human capital that have caused inequality to grow. Future research will hopefully shed further light on this debate.

After having discussed evidence for some general predictions of our framework, next we will present a quantitative version and subsequently use it for a more formal investigation of the quantitative relevance of our mechanism.

2.6 An Enriched Model

In this Section, we extend our model to include heterogeneity in manual (non-learnable) skills and choose some functional forms. We then parameterize our model to reproduce several key stylized facts of the US economy, and use it to evaluate counter-factual policies.

2.6.1 Model Description

A continuum of agents, whose total mass equals one, live for $t \in [0, 1]$, first goes to school until $t = x$ and then works. When in school ($x \leq t$), individuals build learnable human

capital according to the following law of motion:

$$\frac{\partial h_{s,t}}{\partial t} = \beta t^{\beta-1} \alpha_s h_{s,t}. \quad (2.10)$$

Thereafter, $\frac{\partial h_{s,t}}{\partial t} = 0$. This function resembles more conventional human capital functions such as the one due to Ben-Porath (1967), but the time-in-school structure keeps the model computationally simple. Time in school is more productive for the more able and educated, but diminishes over time. $h_{s,0}$ is assumed linear in α_s , so that the two are perfectly correlated. This simplifies the problem significantly at little cost. Non-learnable human capital is given by $h_{m,t} = h_{m,0} = \alpha_m$. Both skills are assumed to be independently drawn from normal distributions (winsorized at three standard deviations from the mean), resulting in a tuple (α_m, α_s) for each individual. When working ($x > t$), individuals derive income from both types of human capital:

$$y_t = w_m h_{m,t} + w_s h_{s,t}. \quad (2.11)$$

From here on out, the individual problem is the same as in equation 2.3 above. We consider overlapping generations such that the population distribution is always in steady state. Let the distribution of type tuples $(\alpha_m, \alpha_s) \in \mathcal{A}$ be denoted by λ . Define human capital aggregates as follows (where $I_{[\cdot]}$ is an indicator function):

$$H_m = \int_0^1 \int_{\mathcal{A}} h_{m,t} I_{[t>x]} d\lambda dt \quad (2.12)$$

$$H_s = \int_0^1 \int_{\mathcal{A}} h_{s,t} I_{[t>x]} d\lambda dt. \quad (2.13)$$

Aggregate production takes place using the following production function:

$$Y = F(H_m, H_s) = A \left[\gamma H_m^\rho + (1 - \gamma) H_s^\rho \right]^{\frac{1}{\rho}}. \quad (2.14)$$

The elasticity of substitution between the two inputs is given by $\frac{1}{1-\rho}$, and γ is a share parameter. We normalize output so that $A = 1$.

A government sets taxes τ_c and $\tau_n(\cdot)$. Its budget is balanced by expenditures G that are assumed not to influence any of the above:

$$\int_0^1 \int_{\mathcal{A}} c_t \tau_c + y_t \tau_n(y_t) I_{[t>x]} d\lambda dt = G. \quad (2.15)$$

Definition 4. *An equilibrium of the model is defined as:*

Wages w_m, w_s ,

allocations H_m, H_s ,

government spending G ,

decision rules for $x, \{c_t\}_{t \in [0,1]} \quad \forall \quad (\alpha_m, \alpha_s) \in \mathcal{A}$

such that given the parameters of the model the following holds:

- individual decision rules solve problem 2.3

- goods markets clear:

$$Y = \int_0^1 \int_{\mathcal{A}} c_t d\lambda dt \quad (2.16)$$

- labor markets clear (equations 2.12 and 2.13)

- wages equal marginal products (of equation 2.14)

- and the government budget constraint is balanced (equation 2.15).

2.6.2 Parameterization

Equilibria of the economy are found numerically. Parameters are set to match moments of the data in the early 2000s. In doing so, the following parameterizations of initial abilities and human capital stocks is used. Let $\tilde{\alpha}_s$ denote a standard normal distribution, winsorized at three standard deviations.

$$\alpha_s = \mu_s + \sigma_s \tilde{\alpha}_s, \quad (2.17)$$

$$h_{s,0} = 1 + (\tilde{\alpha}_s - \underline{\tilde{\alpha}}_s) \psi_s. \quad (2.18)$$

$\underline{\alpha}_s$ is the lowest level of α_s . The lowest level of $h_{s,0}$ is normalized to 1, while average learning ability, the spread in learning ability, and the spread in initial learnable human capital is controlled by parameters. Likewise,

$$h_{m,0} = 1 + (\alpha_m - \underline{\alpha}_m) \psi_m, \quad (2.19)$$

where α_m is standard normal and ψ_m controls the spread of initial non-learnable human capital.

Table 2.4: Parameters and moments

Parameter	Value	Moment	Model	Data
σ	2.857	Elasticity of intertemporal substitution	0.350	0.350
ψ_m	0.141	Earnings variance at start of working life versus overall	0.528	0.500
ψ_s	28.068	Gini coefficient of gross earnings	0.346	0.440
μ_s	0.947	Average share of working age spent in school	0.030	0.034
σ_s	0.225	Variance of share in school	0.002	0.002
β	0.858	Share with zero education after age 18	0.478	0.456
ρ	0.286	Elasticity of substitution in production	1.400	1.400
γ	0.519	Non-learnable share of output	0.248	0.250

Table 2.4 reports data moments. Some of our model parameters are straightforwardly informed by moments of the data, while for others much less clear-cut measures are available. We use the midpoint of the range of elasticities of intertemporal substitution reported in Havranek (2013) to set the same in the model (σ), but that parameter does not influence any of the results we report. The spread of both initial human capitals is important for overall earnings variation, and their relative size helps determine the extent to which that variation is present at age 0. Thus, we target the Gini coefficient of gross earnings as reported by the OECD for the year 2000. We also target a ratio of earnings variance at age 0 versus earnings variance overall of $1/2$. While we do not have a precise estimate for this number from the data, research using the life-cycle of earnings Huggett, Ventura, and Yaron (2011) suggests about two thirds of earnings are pinned down *after* tertiary education. Finally, to determine the average and spread of ability, we target the share of a potential 48 years of working life from age 18 that is spent in school (i.e. college and beyond), the variance of these shares, and the share of pupils who do not spend any time in college. We calculate the data moments from the 2000 Census sample described in the above, where all education beyond 12th grade is counted as taking place during the adult life cycle.

Finally, the parameters in the production function are key to size general equilibrium responses. Unfortunately, no reduced form results on general equilibrium effects between skills as we describe them are available. Instead, we rely on evidence on general equilibrium effects between college educated and non-college educated labor. Here, a large body of evidence suggests an elasticity of substitution of about 1.4 (see for example Katz and Murphy (1992) and Ciccone and Peri (2005)). Because these two groups would both use either type of human capital, we take the view that this is a very conservative estimate of the two elasticity of substitution that is relevant to our model. To tie down the share parameter of the production technology, we target the share of non-learnable human capital in output. Again, no direct evidence is available, so that we tentatively set this target to 25%.

Consumption taxes are set to 7.5%, following the 2003 Figure reported in McDaniel (2007). We estimate the tax function used in the above from tax rates at different levels of average US earnings for 2003, and then do the same for 1983, following Guvenen, Kuruscu, and Ozkan (2014) (we use the same data as those authors). This results in an estimate $\phi_1 = 0.119$ for 2003, which is used for parameterizing the model, and an estimate of $\phi_1 = 0.188$ for 1983, which we use in our counter-factual analysis below. ϕ_0 is set to clear the government's budget constraint.

Table 2.4 also demonstrates the model's ability to match the data. Overall, model moments are close to data moments, although the model does struggle to create sufficient earnings heterogeneity to match the economy's inequality levels.

2.6.3 Results

To analyze the results of tax progressivity, we compare the steady state earnings distribution of the 1983 estimate of ϕ_1 to the steady state distribution with the 2003 estimate. We think of this as a counter-factual reform in which tax progressivity was reduced. The procedure yields a reform that is per definition realistic, both in shape and magnitude. We would not want to argue that our results are empirical in the sense that they have bearing on the change in the period. (For that to be the case, one would want to consider other factors, as well as the transition from one steady state to another.) Rather, we are looking for a counter-factual experiment that gives us a feeling for the effect sizes in our model.

We then turn to measures of inequality. Indeed, reducing the progressivity parameter has increased the 90-10 ratio about one-for-one, which is what we also find in our cross-country analysis. This increase can be almost entirely attributed to the upper half of the distribution i.e. the 90-50 ratio. Again, this is entirely in line with our cross-country findings. These results give us confidence that the model adequately captures the reaction of the earnings distribution to tax progressivity.

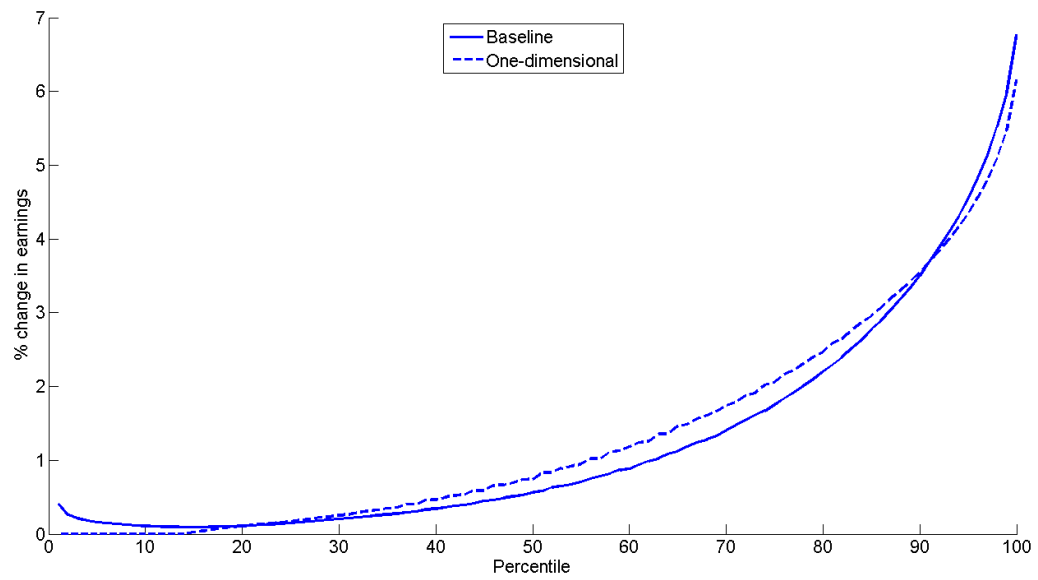
Figure 2.7a shows the results graphically (labeled ‘baseline’). It is apparent that some polarization occurs, but little: the bottom wages grow a few tenths of percent more than those with the lowest wage growth. The top grows by almost 7% more than the lowest point.¹⁰ To bear out polarization given the large increase in inequality in the top half, we show the same graph but restricted to the lower half of percentiles in Figure 2.7b.

There are a number of reasons why one might consider the effect sizes we present conservative. First, the elasticity of substitution between the two skill types may be smaller in practice, leading to larger price effects: the elasticity has been measured in the previous literature using data on college versus non-college educated labor. However, that categorization is a noisy measure of the underlying skills that our theory predicts is relevant. This would lead to an overestimation of the elasticity in a typical regression methodology (e.g. in that of Katz and Murphy (1992)) due to attenuation bias, reducing the price effect (which goes to zero as the elasticity goes to infinity). Second, we have not included leisure, which works as an amplifying mechanism (cf. Guvenen, Kuruscu, and Ozkan (2014)). Third, our view of human capital is a very limited one, because we only focus on time in formal schooling. The same incentives would however also affect learning during the life-cycle, making the overall impact much larger. In addition, in this paper we are focusing on the part of the labor wedge originating exclusively from income taxation. There exist other sources for the labor

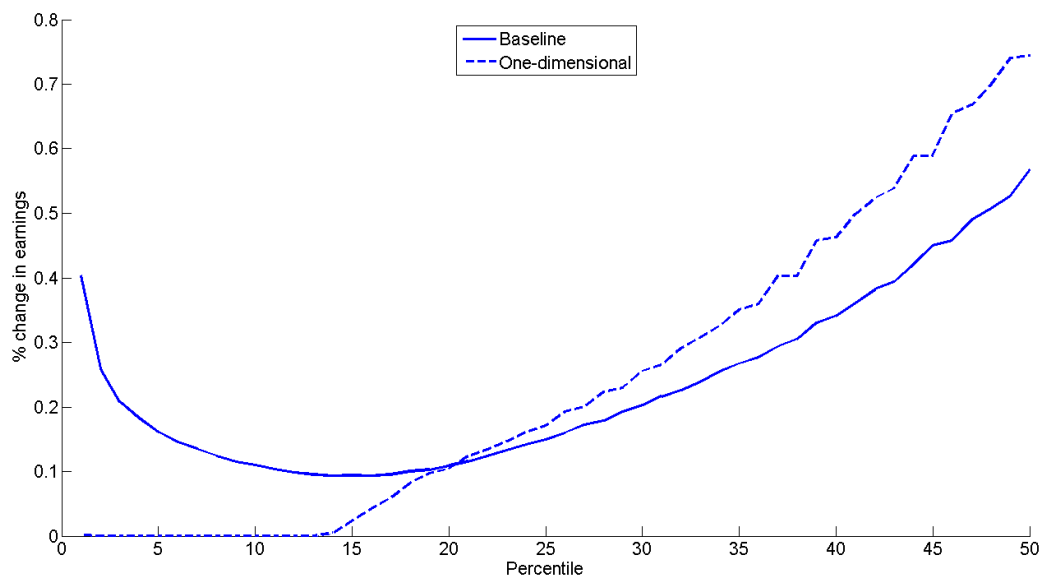
¹⁰For those interested, we report that this is 6% and 45%, respectively, of the equivalent empirical change in the period. As already noted, we do not want to encourage such empirical interpretations too much.

Figure 2.7: Relative earnings change under counter-factual reform

(a) Full distribution



(b) Lower half of the distribution



wedge, in particular discrimination. Since this is outside the current model, we will postpone a detailed discussion of this to the concluding Section 2.7. Finally, potentially also the share of output the model attributes to manual skills, γ , is driving out results. However, we know little about its empirical counterpart - this becomes a suggestion for further research. To investigate the importance of the manual skills share for our results, in the next subsection we will conduct a formal sensitivity analysis of the respective parameter, γ . As will become clear, sensitivity is relatively small. This is reassuring, as it implies that our results are relatively robust to changes in γ .

2.6.4 Sensitivity Analysis

The main moment of which we are uncertain is the one informing γ , the share of output that is contributed by non-learnable skills. At the same time, this parameter is obviously crucial in assessing the importance of our mechanism: in the absence of non-learnable skills output, the model collapses to a uniform human capital model. To make this clear, we re-calibrate the model setting the moment for γ to zero, which results in $\gamma = 0$ (and slight changes to some of the other parameters). Figures 2.7a and 2.7b also show the results in this case (labeled ‘one-dimensional’). While the result is similar for overall inequality, polarization has disappeared. The effect on inequality within the bottom half of the population is now much more straight-forward.¹¹

We provide a more formal analysis of the sensitivity of γ in the remainder of this Section. Our parameters can be interpreted as estimates of an indirect inference procedure: They are the result of minimizing the distance between the data moments described in Table 2.4, the vector of which we will now call \hat{s} , and the model moments that we will call $s(\theta)$ (where θ is the vector of parameters). Defining $\hat{g} = \hat{s} - s(\theta)$, we then used θ to minimize $\hat{g}'I\hat{g}$ (where I is the identity matrix that we use as weights) and reported the argmin $\hat{\theta}$ of our problem in Table 2.4.

Andrews, Gentzkow, and Shapiro (2017) establish a methodology for measuring the sensitivity of parameter estimates to estimation moments. They suggest reporting an estimate of the matrix $\Lambda = -(G'WG)^{-1}G'W$, where G is the Jacobian of the probability limit of \hat{g} at the true parameter values θ_0 , and W is the weighing matrix (the identity matrix in our case). The advantage of their method is that it is computationally simple to find a point estimate of G , and therefore Λ : because our objective vector \hat{g} is additive and only $s(\theta)$ depends on the parameters, we can simply calculate the numeric Jacobian matrix S of our model moments

¹¹The ‘one-dimensional’ graph in Figure 2.7b appears to display a kink that is not actually there: investment in education is always non-zero due to an Inada condition in human capital formation. The visual effect arises because levels of the human capital distribution have been compressed to a percentile scale.

$s(\theta)$ at the estimated parameter value $\hat{\theta}$. In short, we have that our *sensitivity* estimate is given by $\Lambda = S^{-1}$.

How should these sensitivity estimates be interpreted? Entry λ_{ij} of Λ tells us, roughly, how large the local impact of a change in data moment j is on parameter i . It can be used to calculate the asymptotic bias in our estimates associated with an alternative hypothesis on the data moments, as long as the alternative is sufficiently close to the data moments we report. More straightforwardly, it can be used to verbally discuss the sensitivity of our estimates to the data moments. That is a particularly appealing feature in light of the uncertainty around some of the data moments that we report above. Because a unit change in the data moments is not always easy to interpret, we instead opt to report results relevant to a 1% change of each data moment. This is achieved by multiplying λ_{ij} by a percent of data moment j . The results are in Table 2.5.

Table 2.5: Sensitivity Analysis

Moment nr.	1	2	3	4	5	6	7	8
σ	-0.23	0.00	0.00	0.00	0.00	0.00	0.00	0.00
ψ_m	0.00	-0.01	-0.03	-0.08	0.51	-0.00	-0.00	-0.01
ψ_s	0.00	50.95	335.41	-566.61	1901.24	-23.21	-2.41	129.65
μ_s	0.00	0.05	-0.02	0.29	4.45	0.01	0.00	0.02
σ_s	0.00	-0.00	0.00	-0.03	-0.19	-0.00	-0.00	-0.00
β	0.00	0.02	0.00	0.05	1.60	0.00	0.00	0.01
ρ	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
γ	0.00	0.01	0.03	0.02	-0.20	0.00	0.00	0.02

Note: Model moments are 1 – Elasticity of inter-temporal substitution, 2 – Earnings variance at start of working life versus overall, 3 – Gini coefficient of gross earnings, 4 – Average share of working age spent in school, 5 – Variance of share in school, 6 – Share with zero education after age 18, 7 – Elasticity of substitution in production, 8 – Manual human capital share of output. See also Table 2.4.

Two parameters, σ and ρ , are only sensitive to the one moment on which they depend by a closed-form relation (the latter's sensitivity measure is zero in the table due to round-off). ψ_s takes on larger values, and so is generally more responsive in level terms. ψ_m and ψ_s react most heavily to moments that describe the distribution of earnings and schooling. As we would hope, parameters describing learning ability and the formation of human capital indeed react most strongly to those moments that describe the distribution of schooling. The parameter γ reacts strongly to the 5th moment, the variance of schooling, which clearly

plays an important role in the determination of the model's parameters.

As discussed above, we have very little information about the 'non-learnable share of output', γ , which is the eighth and last data moment in Table 2.5 above. It turns out that this parameter does play some role in the determination of γ , albeit not a large one. That there is some sensitivity is quite in line with our expectation, given the analysis included above where we set $\gamma = 0$. The fact that the sensitivity is not extremely large is reassuring, since it implies that our results would relatively little if our target of γ was somewhat off. We remain with the conclusion that the importance of manual skills in the overall economy is an important determinant of the strength of our mechanism, but that we unfortunately know little about it.

2.7 Conclusion

This paper has analyzed the effect of tax incentives on cognitive skills, in a model where (learnable) cognitive and (non-learnable) manual skills jointly produce earnings. It has also attempted to argue why this is a relevant view of the labor market, combining general equilibrium elements from the literature on skill-biased technological change and incentive elements from the literature on human capital formation. In doing so, it has provided an alternative mechanism through which labor market polarization may arise.

In the paper we focus exclusively on the part of the labor wedge originating from taxation. An important additional source of the labor wedge originates from discrimination. Over the second half of the 20th century (labor market) discrimination against women and non-white groups arguably decreased a lot. There is growing evidence that the decline in discrimination has been quantitatively important for US macroeconomic outcomes. Dwyer (2013) provides evidence that polarization in employment has been driven to a substantial part by women increasingly entering the labor market, primarily in the tails of the distribution. Hsieh et al. (2016) estimate that about 25% of US output per capita growth between 1960 and 2010 can be attributed to an improved allocation of talent due declines in discrimination in the labor market and in access to education. Decreasing the price of education for a substantial share of the working population would have a similar effect as the decline in tax progressivity, by increasing the relative payoff of spending time in school. Similarly, if declines in discrimination take place in the form of 'breaking the glass ceiling', they might over-proportionally improve labor market outcomes for high-earning women, again resembling declines in tax progressivity. Potentially, these results therefore imply that the decline of progressivity of the effective labor wedge has been a lot larger than the decline in the explicit tax wedge. In this case, our results present a definitive lower bound on the supply-side polarization channel discussed here.

Future research may lead in a number of directions. First, fundamental questions on our model of the labor market remain of interest. For example, credibly exogenous variation in skill levels might illuminate the prices paid for different levels (or bundles) of skills. Second, further research into the distributional effects of reduced discrimination against minority groups in the labor market seems warranted. Finally, while the emphasis in this paper has been on positive implications, one might ask what optimal tax and education policies look like in a model like ours. In the presence of general equilibrium effects, tax disincentives to the formation of human capital are more harmful than is traditionally assumed, likely warranting less progressive tax schedules.

2.8 Appendix

2.8.1 PCA Results

The table below displays the full PCA results. Each column represents the correlation between a component and the original variables. The table begins with the component that explains the largest share of variance, then the second largest, and so forth.

Data	0.49	-0.31	0.25	0.51	0.49	0.12	0.13	0.31	0.16	0.42	0.60	0.62	0.49	-0.31	0.77	-0.06	0.81	0.70
People	0.45	-0.27	0.34	0.39	0.56	0.47	0.15	0.05	0.19	0.20	0.57	0.68	0.70	0.32	0.70	0.33	0.46	0.36
Things	-0.34	0.12	-0.49	-0.15	-0.38	-0.42	-0.15	0.30	0.56	-0.19	-0.01	-0.45	-0.55	0.11	-0.29	0.09	0.47	0.45
GED Reasoning	0.47	-0.36	0.22	0.50	0.62	0.27	0.09	0.33	0.16	0.38	0.64	0.68	0.52	0.01	0.66	-0.19	0.86	0.77
GED Mathematical	0.46	-0.39	0.22	0.49	0.45	0.04	0.06	0.43	-0.02	0.55	0.56	0.62	0.45	-0.01	0.64	-0.35	0.83	0.72
GED Language	0.48	-0.38	0.27	0.54	0.61	0.34	0.04	0.43	0.14	0.32	0.65	0.69	0.65	0.04	0.67	-0.22	0.75	0.70
Specific Vocational Preparation	0.36	-0.22	0.04	0.46	0.36	0.03	0.17	0.24	0.29	0.31	0.46	0.61	0.31	-0.20	0.58	-0.13	0.88	0.75
Intelligence	0.51	-0.44	0.26	0.54	0.60	0.25	0.05	0.36	-0.02	0.39	0.72	0.69	0.52	-0.09	0.54	0.02	0.80	0.75
Verbal	0.51	-0.49	0.37	0.61	0.55	0.34	0.09	0.19	0.00	0.37	0.76	0.68	0.67	-0.15	0.60	-0.01	0.72	0.69
Numerical	0.43	-0.62	0.31	0.38	0.59	0.02	-0.06	0.38	0.11	0.59	0.58	0.53	0.36	-0.02	0.53	-0.22	0.71	0.72
Spatial	0.05	0.41	-0.59	0.07	0.13	-0.63	0.21	0.12	0.19	0.12	-0.15	0.02	-0.19	0.16	0.14	-0.03	0.50	0.56
Form Perception	0.05	-0.09	-0.32	0.49	-0.04	-0.09	-0.06	0.34	0.42	0.30	0.04	-0.11	-0.00	-0.03	-0.04	0.21	0.63	0.76
Clenical Perception	0.53	-0.51	0.60	0.42	0.38	0.45	-0.45	0.34	0.07	0.41	0.73	0.40	0.41	-0.09	0.41	-0.04	0.47	0.56
Motor Coordination	-0.44	-0.07	-0.47	-0.37	-0.55	0.01	-0.35	0.05	0.25	-0.14	0.08	-0.51	-0.24	0.08	-0.42	0.28	0.27	0.49
Finger Dexterity	-0.63	-0.14	-0.31	-0.00	-0.37	-0.05	-0.40	0.15	0.39	-0.03	0.04	-0.48	-0.15	0.08	-0.36	0.22	0.40	0.57
Manual Dexterity	-0.57	0.21	-0.54	-0.38	-0.45	-0.48	-0.13	-0.19	0.22	-0.47	-0.57	-0.70	-0.36	0.24	-0.53	0.11	0.15	0.15
Eye-Hand-Foot Coordination	-0.14	0.44	-0.51	-0.43	-0.30	-0.26	0.48	-0.20	-0.22	-0.69	-0.38	-0.17	-0.61	0.46	0.09	0.20	-0.16	0.01
Color Discrimination	-0.28	0.46	-0.05	-0.25	-0.09	0.16	0.51	0.60	0.39	0.00	-0.24	-0.27	-0.14	-0.00	-0.02	0.28	0.20	0.43

Chapter 3

Economics PhD Programs in Europe: Completion Times and More

3.1 Introduction

Recent years have seen an increasing convergence of economics PhD programs in Europe towards their US counterparts. Increasingly, a number of European programs offer structured programs that include a significant coursework component, and are able to provide funding for the entire program length. While this convergence has been noted, little structured information is available on these European programs. We take a first step towards filling this gap.

Using a unique hand-collected dataset, we add to results by Stock et al. (2009) and Stock and Siegfried (2014) on completion times in US graduate programs in economics. We find that their conclusions on completion time hold in European structured graduate programs as well. In the past 5 years, median completion times have been rising steadily, and the median is now approaching 6 years. We also provide some qualitative description of the differences between European programs and their US counterparts, in particular with regards to program structure and funding.¹ Because our analysis focuses on subset of all European programs, with emphasis on including good programs that resemble top US programs in structure, our results should not be taken as representative of European economics PhD education more

¹This project started as an internal project at the European University Institute (EUI), where we tried to understand the differences between the EUI Economics Doctoral Program and its closest competitors, in particular with regards to completion time. While such information was readily available for US programs, we had to hand-collect information on European programs by going through the CVs of over 700 recent job market candidates. Subsequently, we realized that our findings could benefit the economics community at large, and therefore merit wider dissemination.

broadly.

Our analysis focuses on job market candidates. These are the students that compete on the international academic job market for economics graduates.² We believe this to be the policy-relevant subsample for three reasons. First, delivering internationally competitive candidates is the stated goal of many European economics PhD programs. That job market is the highest international standard for young researchers in economics. Second, this measure ensures that the students we consider are at the same stage of their career as their US counterparts. Third, over the last years, the international job market has become an important recruiting device for economics PhD graduates outside the narrowly defined academic sector. Increasingly, international organizations, such as central banks, governments and also the private sector are hiring through the economics job market.

The paper also investigates how completion time and initial placement quality vary with personal researcher characteristics. For each candidate in the sample, we collect data on gender, field, field of undergraduate studies and their initial placement. Candidates who go on the market in the sixth year of their PhD have a significantly higher probability of top ranked initial placements. We find field and gender do not significantly correlate with completion time nor with placement quality. We do find however, that the probability of placing in a top ranked institution is significantly lower for economics PhD candidates trained in social sciences, compared to candidates trained in economics, finance, business, engineering or natural sciences. We interpret this finding as evidence for the importance of formal mathematical training for the successful completion of an economics PhD.

3.2 Economics PhD Programs in Europe

Selection of Programs We established a list of top European programs from a variety of sources. In particular, we emphasized that they should have international recruiting and placement, as well as a structured graduate program (including coursework), which makes them comparable to top US programs in style. Several publicly available rankings aided our search for candidate programs. Our procedure resulted in a list of 21 programs. Of these, 5 are located in the United Kingdom (Cambridge, LSE, Oxford, UCL, Warwick), 4 in Spain (Autonoma Barcelona, Carlos III, CEMFI, UPF), 3 in Germany (Bonn, Frankfurt, Mannheim), 2 in France (Paris School of Economics, Toulouse), 2 in Italy (Bocconi, EUI), 2 in the Netherlands (Tilburg, Tinbergen Institute), 2 in Sweden (IIES, Stockholm School of Economics), and 1 in Switzerland (Zurich). We do not purport to establish that these are the 21 best programs in Europe, nor would we want to rank them. But the list does include

²For more information on the economics job market, visit <https://www.econjobmarket.org/index.php>. The organization describes itself as a "non-profit clearinghouse for applications to PhD level jobs in economics."

all usual suspects for the top spots, so that we would certainly expect that, for example, the top 10 programs (whichever they are) are included in this set.

Program Structure Table 3.1 provides some information on program structure for each of the European programs. Generally, the setup of these programs is very similar to that of US programs: they consist of a coursework phase, where the first year consists of core courses and the second year consists of electives and moves students towards the research frontier. The balance between courses and initial research in the second year varies from program to program, while the first year programs are largely standard and very similar to those in US programs.

There is one big difference in setup to US programs: in many of the European programs, the coursework phase is treated as a separate degree program, and leads to titles such as MPhil or MRes. Thus, in some of these programs PhD students are just a subset of a masters degree class for the first one or two years (for example CEMFI), while others are entirely integrated as in a US-style system (for example EUI). Further detail is provided in Table 3.1.

In all programs under consideration, coursework is targeted at future PhD researchers. Why then offer this coursework for a terminal masters degree as well? We suspect that this is largely due to the current setup of European academia and the structure of public funding for higher education. Following the Bologna process, the typical European student completes a 3-year Bachelors degree in his field of interest, followed by a 1- or 2-year Masters degree. State funding (most of the programs are housed by public institutions) is often structured the same way, where it is beneficial to have class sizes above the typical number of students in a PhD program. Thus, economics departments fit the 'ideal' US type program into the European system by making the coursework phase a Masters degree. In addition, some programs use this to select the best students from a large pool that undergoes initial coursework (for example UPF). The demands that first-year coursework places on students leads to the admission of mostly (or in some cases exclusively) students who already completed a 'regular' Master's degree elsewhere, sometimes in the same university.

One possibility that the above structure raises is that students might switch programs when they enter the research phase. We find that such switching is exceedingly rare. While it occurs that students take Masters degree in one of these programs and then enter another, they then typically retake the entire coursework phase of their new program. Summarizing all of this, we concluded that these two-step programs can safely be considered PhD programs with a coursework component, similar to their US counterparts.

Program Funding There are large differences in funding from program to program as well. Not too dissimilar from the US, all of these programs can in principle fund students through the entire length of a PhD. However, in some cases funding is insecure from year to year, or only provided in return for teaching and research assistantships. Finally, there are large differences in the availability of funding across any one cohort, with some programs providing funding to all admitted (for example SSE) and others separating the admission decision from the funding decision entirely (for example Cambridge and Oxford).

An important difference to fully integrated US PhD programs arises in the European programs that separate their course phase into a stand-alone master degree. Here, the majority of the programs do not provide funding for the master stage. Funding is then restricted to the research phase of the PhD, following the initial one or two years of coursework.

3.3 Data Collection and Processing

Data collection proceeded in two stages. In the first stage, we looked up current job market candidates for each program (this was in 2015-2016) and placement results of past job markets (2011-2012 through 2014-2015) online, whenever data were available. We searched for each candidate's CV using personal websites, professional websites, and LinkedIn profiles. This led to some, but very limited, missing data for students who are listed as job market candidates but whose entry dates cannot be established.

In the second stage (in the Summer of 2017), we collected job market outcomes and additional covariates per candidate. Job market outcomes were codified for 'quality' as follows. Outcomes were classified in three different classes, academic, institutional or private sector. Within each class of placements, we assigned specific institutions as member of the Top, Middle or Low group of institutions within that class. For academic institutions, we made use of the IDEAS/REPEC ranking of Top Economics Institutions, as of June 2017, to assign universities to groups.³ For institutional jobs, we ranked prominent international institutions such as IMF, ECB or Worldbank, and top national institutions of large countries (Fed, Banque de France etc.) as Top, less prominent international institutions as Middle and national institutions of smaller countries as Low. For private sector jobs, our decision rule was based on international reputation of the company, without resorting to a formal criterion. While the IDEAS/REPEC ranking is just one of many possible academic rankings, it is easily accessible, computed based on transparent rules and comprehensive in coverage. In addition to the institutional quality ranking just discussed, we created a job quality ranking depending on the job title of the first placement job.

³The current ranking can be accessed via <https://ideas.repec.org/top/top.inst.all.html>. The ranking is updated continuously, so the current rankings might differ slightly from the ones we used for our analysis.

Additional covariates collected include gender, field (codified as micro, macro, applied, econometrics, or finance), fields of undergraduate studies (economics, business, natural sciences/engineering, social sciences/humanities), age, and nationality. (Data on the latter two turned out to be missing in many cases, so that we did not use them for analysis.) We also returned to our original sources and compared reported placements for 2016 versus the list of names that was reported to be on the market in 2015-2016. This gives us a clear idea of the extent to which collecting reported past candidates is representative of actual past candidates. We discuss these results below.

As we explain above, we focus on candidates that enter the international job market at the end of their PhD. Delivering competitive candidates to this job market is the stated goal of most programs in our sample. Our approach makes our data more reliable and more comparable to US data on completion times than European data from official sources would be. European programs differ strongly in how they account for coursework time and time spent on the job market (which is sometimes after defense), as we described above. Therefore, a detail-oriented approach and knowledge of the program structure is necessary to attribute the right de-facto number of years to the time spend in a PhD program.

We determined each candidate's entry year into the program. Our measure for completion time, time to job market, is the difference between the job market year and the entry year. We include the relevant course period and the time spent on the job market. The measure is not dependent on the timing of the defense.

3.4 Sample Selection and Representativeness

The attempt of this paper is, in first instance, to compare European job market candidates to their North American counterparts. The main selection issue that our method of data collection faces is whether reported past job market candidates (placements) correspond to actual past job market placements. Here, two issues might occur. First, candidates might be added to placement lists when they get jobs outside of the academic job market. Second, candidates who do not place well or not at all may not be reported among past placements.

As mentioned above, our two-staged data collection approach allows us to examine this issue for the 2015-2016 job market. We found the former issue to be minor, and therefore do not report it. The latter issue is somewhat more prevalent, although only in some programs. We report the results in Table 3.2 below. In one case (UPF) a university stopped reporting placements altogether. For a couple, job market candidates were not yet listed during our first stage of data collection.

This issue, while relevant to prospective students, is most relevant to our research setup if it

influences our results on completion time significantly. It turns out that the issue is minor: those who are listed as placements took on average 6.05 years by our measure, while those that were not reported as placements took 5.90 years.

Another issue is missing data. For few programs, one or several years of job market data are missing. Typically these are recent (Cambridge, Warwick, and Zurich) or older years (UCL, UPF, Warwick), or years for which a smaller program did not deliver job market candidates (CEMFI). However, we do not believe that this influences our conclusions with regards to completion time.

Importantly, we do not claim that our sample is a representation of entering students: Some may drop out or be asked to leave at various stages. Some students may not enter the job market and therefore go unlisted on websites. For the EUI, we have access to administrative data on all PhDs awarded. On average, about half of the PhD recipients go on the Academic Job Market and are publicly listed as such. For the years 2012 to 2015, the average job market candidate took a quarter of a year longer to obtain their degree than the average degree recipient. While these are interesting statistics, they are not relevant to our research setup.

3.5 Results

Completion Times Tables 3.3, 3.4, 3.5 and 3.6 below contain our results on completion times. Average and median completion times have been rising since 2013. Both the average (Table 3.3) and median (Table 3.4) are at 6 years for the 2016 job market cohorts. This finding is remarkably consistent across programs, with both averages and medians lying between 5 and 7 years for all programs. We miss information for a few years for some programs (whenever the reported number of observations is zero in Table 3.5), but our overall number of observations is large at 736. We do not observe qualitative changes to our results when calculating observation-weighted averages (Table 3.6).

To provide some further insights, Figure 3.1 provides a histogram of completion times across all programs for the last two years in our sample. Completion times are rather concentrated around 6 years. While 5 and 7 years of completion time also occur frequently, almost no students finish in 4 years or less.

Covariates of Completion Times We report results from an ordered probit model regressing Time to Job Market (completion time) on Year of Job Market, Field of Economics, Gender, PhD Institution and Field of Undergraduate Studies, estimated with robust standard errors. For each covariate presented, we report the probability of observing a certain

Time to Job Market for each level of the covariates, holding all other covariates at their respective sample means. Confidence intervals shown are 95 percent confidence intervals.

The probabilities of observing particular completion times for each year of the sample are shown in Figure 3.2. This confirms the raw completion times results reported in Tables 3.3 through 3.6. Completion in 6 or 7 years becomes significantly more likely over the sample period, with 6 years being significantly more likely in 2015 and 2016, the opposite of what is observed in the early sample years. Going on the job market in the 4th year of the PhD, while still likely in 2009 and 2010, is becoming increasingly unlikely towards the end of the sample.

Figure 3.3 shows that completion times of female candidates are slightly longer than their male counterparts, however the differences are not statistically significant. Figure 3.4 shows that the same is true for PhD candidates in different subfields of economics.

Covariates of Placement Quality We estimate a similar model to the one above, this time using our placement quality index as dependent variable, pooled over all classes of placement. In addition to the covariates reported above, we also include Time to Job Market as an additional explanatory variable. An important qualification to stress is that this does not allow us to infer any causal link from completion time to placement quality and that the reported association is purely statistical in nature. Results reported are obtained in the same way as in the previous subsection.

Figure 3.5 shows probabilities of placing in a top, middle or low ranked job within the sample graduating in 4, 5, 6, 7 or 8 years. While for a duration of 4 or 5 years, all placement qualities are statistically equally likely, going on the job market after 6 years is associated with a significantly higher probability of placing in a top ranked job relative to a middle or low ranked one. This effect attenuates again for graduates with a duration above 6 years. Figure 3.6 plots the average completion time and placement quality for graduates of each PhD-granting institution in the sample. As is apparent from the figure, average completion time and average placement quality are associated positively in the sample. While, as already mentioned above, this does not suggest that longer completion times are causing better placement, it is a clear indicator that high quality candidates are taking additional time to go on the job market. Since placing their candidates well is the stated goal of the departments in our sample, this has implications for the desired funding structure of their PhD programs.

Figure 3.7 shows that Gender does not vary significantly with placement quality. Since comparing placement quality by subfield of economics is not particularly meaningful, we instead report placement quality by undergraduate background of PhD candidates in Figure

3.8. While there is no statistical difference between candidates that have a background in Economics, Business, Natural Sciences or Engineering, candidates with a social science or humanities background place significantly worse relative to the other backgrounds. We view this as evidence for the importance of sufficient formal training.

Assistant Professor Subsample Since the primary goal of many PhD programs is preparing candidates for a career in Academics, we present some additional results on the subsample of candidates with a first placement job title as 'Assistant Professor'.

Sample frequencies for the completion times of this subsample are presented in Figure 3.9. The relative frequency of 6 vs 5 years of completion times is skewed in favor of 6 years, with 29 vs 46 percent. This compares to relative frequencies of 32 vs 39 percent of the entire sample. Figure 3.10 shows that placing at a top-ranked university is unlikely for all durations, but point estimates are slightly increasing from 5 to 7 years of time to completion.⁴ Comparing top placement probability changes across subsamples of different duration reveals insignificant estimates (Figure 3.11). However, when plotting average placement institution's quality and average completion times for the assistant professor subsample, there is again a positive correlation. This relationship is plotted in Figure 3.12.

Taken together, results the subsample of 'Assistant Professor' confirm the impression of the full sample analysis: While the results and estimates presented do not allow for causal interpretation, statistically completion times and placement quality are positively related.

3.6 Conclusion

Recent years have seen an increasing convergence of economics PhD Programs in Europe to their US counterparts. Completion times in the top programs have steadily risen to, and now reached, a median of 6 years. This brings them rather close to completion times in US programs as surveyed by Stock et al. (2009) and Stock and Siegfried (2014). However, program and funding structures remain different due to institutional factors. Our findings may therefore be of relevance to funding authorities and administrators.

Our results suggest that higher placement quality is statistically associated with longer completion times. In addition, we uncover a number of facts relating completion times and placement quality to personal researcher characteristics. The data do not show evidence of systematic differences by gender in either the duration or the placement quality of European economics PhD programs. On the other hand, undergraduate background turns out to

⁴Only 23 candidates place as Assistant Professors after completing their PhD in 4 or less years, which makes interpretation of the estimated probabilities for those subsamples difficult.

be a significant predictor of success in an economics PhD program, if measured by initial placement quality.

3.7 Appendix

3.7.1 Tables and Figures

Table 3.1: Program Details

UAB	The Graduate Program consists of a two-year Master in Economic Analysis, followed by a research phase. The entire path is counted.
Bocconi	The program is officially described as a four year program, with the first two years dedicated to course work. However, our analysis shows that candidates take substantially longer to finish their PhD.
Bonn	The program is structured into 4 years, of which the first two are mainly course work. Time can be extended into the fifth year for the academic job market.
Cambridge	The full program is split into a one-year MPhil (coursework) phase and a PhD phase (research). Both are counted as time to completion.
Carlos III	The Graduate Program consists of a two-year Master in Economic Analysis, followed by a three year PhD in Economics. Sufficient performance of the former provides entry to the latter. The entire path is counted.
CEMFI	The PhD program starts out with two years of coursework, which is taken jointly with a master's program. Some master's students subsequently enroll as PhD students. In either case, the entire path is counted.
EUI	Program is entirely standardized, with coursework as part of the PhD program. There was a small terminal master's program in the past consisting of part of the same coursework.
Frankfurt	The program is officially described as a four year program, with the first two years dedicated to course work. However, our analysis shows that candidates take substantially longer to finish their PhD.
IIES	The PhD program is organized jointly with the Department of Economics of the University of Stockholm. Entry into IIES is competitive out of the program. We count the full time spent in the PhD program, also if part of it was spent outside of the IIES.
LSE	The full program is split into a two-year MRes (coursework) phase and a PhD phase (research). Some who obtained a previous master's degree (usually a terminal MSc from LSE) may be allowed to complete the MRes in one year instead of two. Both the time spent on the MRes and the PhD phase are counted as time to completion, but previous degrees are not.
Mannheim	The program lasts 5 years, of which the first two years are course work. Funding is committed for the entire period.

Oxford	The full program is split into a two-year Mphil (coursework) phase and a DPhil phase (research). Both are counted as time to completion.
Paris SE	The Paris School of Economics (PSE) is a collection of Economics departments in Paris. PhD candidates from these schools are listed as PhD candidates of PSE. Three different subsets of this set of schools each jointly offer a master's degree, which consists of one year of core coursework (the Master 1) and one year of advanced coursework (the Master 2). Subsequently, students may be admitted to a PhD program. The entire path is counted.
SSE	The program starts with a two year course phase after which two to three years of research follow.
Tilburg	The five-year graduate program consists of a two-year Research Master and a three-year PhD program. The entire path is counted.
Tinbergen	The Tinbergen Institute is a joint graduate school and research institute of the Erasmus University Rotterdam, the University of Amsterdam, and the Free University Amsterdam, Netherlands. Tinbergen offers a two-year MPhil degree, after which students can be offered doctoral positions in one of the three universities. The three universities also hire PhD students for four year positions that do not require formal coursework (further to previous degrees), and some of these are listed as Tinbergen placements. Because we cannot distinguish between the two, we list all students that the Tinbergen Institute lists and count time spent in the MPhil as well.
Toulouse	The TSE doctoral program consists of a Master 2 (French university system) in Econometric Theory and Econometrics, which is explicitly part of the 'doctoral track', a DEEQA degree, which is essentially the second year of coursework, and a research phase. The entire path is counted.
UCL	The program is structured into MRes. (one year, coursework), MPhil. (second year, research) and PhD (following two years). Thereafter, students have another year to complete their thesis with full student status.
UPF	The typical path towards a PhD at UPF includes one year of core courses in an MSc program, one year of advanced courses in an Mphil program, and then a research phase. While the MSc is also a large terminal degree (at least with respect to UPF), it is part of the core sequence of courses for a UPF PhD. Thus, the entire path is counted.
Warwick	The program is structured into a two-year MRes. followed by a 4 year PhD (total: 2+4). Students should submit towards the end of year 3 of the PhD and go on the job market in year 4.

Zurich The program has a two year course phase followed by a research phase which is not formally structured.

Table 3.2: Percentage of reported job market candidates eventually listed as placements for the Academic year 2015/16

Percentage listed	2016
EUI	100%
LSE	100%
Oxford	43%
Cambridge	
UPF	0%
Carlos III	100%
Toulouse	100%
Paris SE	100%
Tinbergen	75%
Tilburg	13%
Autonoma Barcelona	86%
CEMFI	100%
UCL	83%
Warwick	
Zurich	
Bonn	89%
Mannheim	92%
IIES	100%
SSE	25%
Bocconi	100%
Frankfurt	83%
Average	77%

Table 3.3: Average Time to Job Market (years)

Program	2012	2013	2014	2015	2016	Average
Autonoma Barcelona	6.00	6.20	5.50	6.00	6.14	<i>5.97</i>
Bocconi	6.50	5.88	6.20	6.86	5.80	<i>6.25</i>
Bonn	4.80	5.38	5.67	4.75	5.56	<i>5.23</i>
Cambridge	4.80	5.33	5.67			<i>5.27</i>
Carlos III	5.89	6.29	5.67	5.56	5.50	<i>5.78</i>
CEMFI	5.75		6.50	6.20	6.00	<i>6.11</i>
EUI	4.45	4.75	5.18	4.89	5.10	<i>4.88</i>
Frankfurt	5.44	5.60	5.78	5.75	6.00	<i>5.71</i>
IIES	6.00	6.00	6.00	7.00	6.60	<i>6.32</i>
LSE	6.31	5.79	6.79	6.14	6.30	<i>6.27</i>
Mannheim	5.60	5.00	5.38	5.78	5.75	<i>5.50</i>
Oxford	6.14	5.31	5.50	5.56	6.14	<i>5.73</i>
Paris SE	6.13	6.00	6.00	6.27	6.30	<i>6.14</i>
SSE	6.00	5.50	5.50	5.33	6.00	<i>5.67</i>
Tilburg		4.80			5.50	<i>5.15</i>
Tinbergen	5.32	4.93	5.44	5.39	5.94	<i>5.40</i>
Toulouse	5.70	5.50	5.75	6.60	6.80	<i>6.07</i>
UCL			6.56	6.29	6.42	<i>6.42</i>
UPF				5.50	6.50	6.00
Warwick		5.64	5.50	6.00		<i>5.71</i>
Zurich	4.00	5.00	5.20	7.50		<i>5.43</i>
<i>Average</i>	<i>5.58</i>	<i>5.49</i>	<i>5.78</i>	<i>5.97</i>	<i>6.02</i>	<i>5.76</i>

Table 3.4: Median Time to Job Market (years)

Program	2012	2013	2014	2015	2016	Average
Autonoma Barcelona	6	6	6	6	6	<i>5.90</i>
Bocconi	7	6	6	6	6	<i>6.10</i>
Bonn	5	5	5	5	6	<i>5.20</i>
Cambridge	4	5	6			<i>5.00</i>
Carlos III	6	6	6	6	6	<i>5.90</i>
CEMFI	6		7	6	6	<i>6.13</i>
EUI	4	5	5	5	5	<i>4.80</i>
Frankfurt	6	6	6	6	6	<i>5.90</i>
IIES	6	6	6	7	7	<i>6.40</i>
LSE	6	6	7	6	6	<i>6.20</i>
Mannheim	5	5	5	6	6	<i>5.40</i>
Oxford	6	5	6	5	6	<i>5.60</i>
Paris SE	6	6	6	6	6	<i>6.00</i>
SSE	6	5	5	5	6	<i>5.40</i>
Tilburg		5			6	<i>5.50</i>
Tinbergen	5	5	5	5	6	<i>5.20</i>
Toulouse	6	6	6	7	7	<i>6.10</i>
UCL			6	6	7	<i>6.17</i>
UPF				6	6	<i>5.75</i>
Warwick		6	6	6		<i>5.83</i>
Zurich	4	5	5	8		<i>5.38</i>
<i>Average</i>	<i>5.50</i>	<i>5.47</i>	<i>5.68</i>	<i>5.84</i>	<i>6.06</i>	<i>5.71</i>

Table 3.5: Number of observations

Program	2012	2013	2014	2015	2016	Total
Autonoma Barcelona	7	10	4	2	7	30
Bocconi	4	8	5	7	5	29
Bonn	5	8	9	4	9	35
Cambridge	5	6	3	0	0	14
Carlos III	9	7	9	9	4	38
CEMFI	4	0	2	5	1	12
EUI	11	12	11	9	10	53
Frankfurt	9	5	9	8	6	37
IIES	1	7	1	1	5	15
LSE	16	14	14	21	10	75
Mannheim	5	5	13	9	12	44
Oxford	7	16	6	9	7	45
Paris SE	8	9	6	11	10	44
SSE	5	6	6	3	4	24
Tilburg	0	5	0	0	8	13
Tinbergen	25	15	25	31	16	112
Toulouse	10	2	8	10	10	40
UCL	0	0	9	7	12	28
UPF	0	0	0	10	6	16
Warwick	0	11	6	6	0	23
Zurich	1	1	5	2	0	9
<i>Total</i>	<i>132</i>	<i>147</i>	<i>151</i>	<i>164</i>	<i>142</i>	<i>736</i>

Table 3.6: Observation Weighted Average Time to Job Market

Program	2012	2013	2014	2015	2016	Average
Autonoma Barcelona	6.00	6.20	5.50	6.00	6.14	6.03
Bocconi	6.50	5.88	6.20	6.86	5.80	6.24
Bonn	4.80	5.38	5.67	4.75	5.56	5.34
Cambridge	4.80	5.33	5.67			5.21
Carlos III	5.89	6.29	5.67	5.56	5.50	5.79
CEMFI	5.75		6.50	6.20	6.00	6.08
EUI	4.45	4.75	5.18	4.89	5.10	4.87
Frankfurt	5.44	5.60	5.78	5.75	6.00	5.70
IIES	6.00	6.00	6.00	7.00	6.60	6.27
LSE	6.31	5.79	6.79	6.14	6.30	6.25
Mannheim	5.60	5.00	5.38	5.78	5.75	5.55
Oxford	6.14	5.31	5.50	5.56	6.14	5.64
Paris SE	6.13	6.00	6.00	6.27	6.30	6.16
SSE	6.00	5.50	5.50	5.33	6.00	5.67
Tilburg		4.80			5.50	5.23
Tinbergen	5.32	4.93	5.44	5.39	5.94	5.40
Toulouse	5.70	5.50	5.75	6.60	6.80	6.20
UCL			6.56	6.29	6.42	6.43
UPF				5.50	6.50	5.88
Warwick		5.64	5.50	6.00		5.70
Zurich	4.00	5.00	5.20	7.50		5.56
<i>Average</i>	<i>5.61</i>	<i>5.50</i>	<i>5.74</i>	<i>5.83</i>	<i>6.02</i>	<i>5.77</i>

Figure 3.1

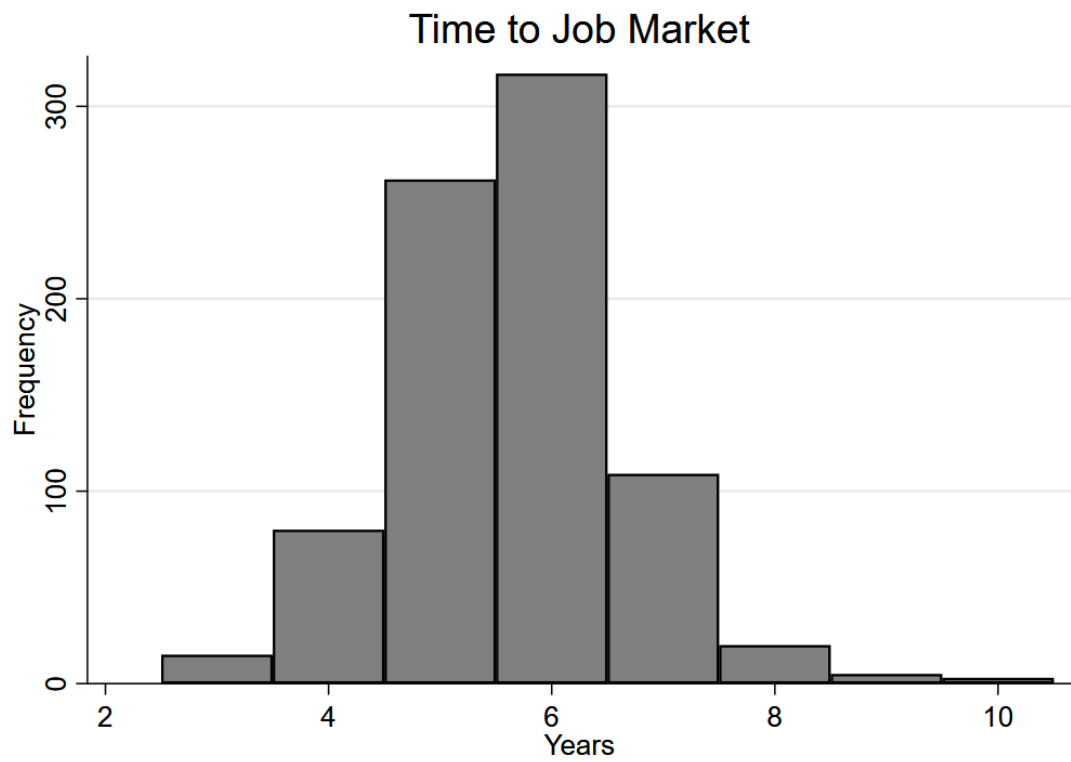


Figure 3.2

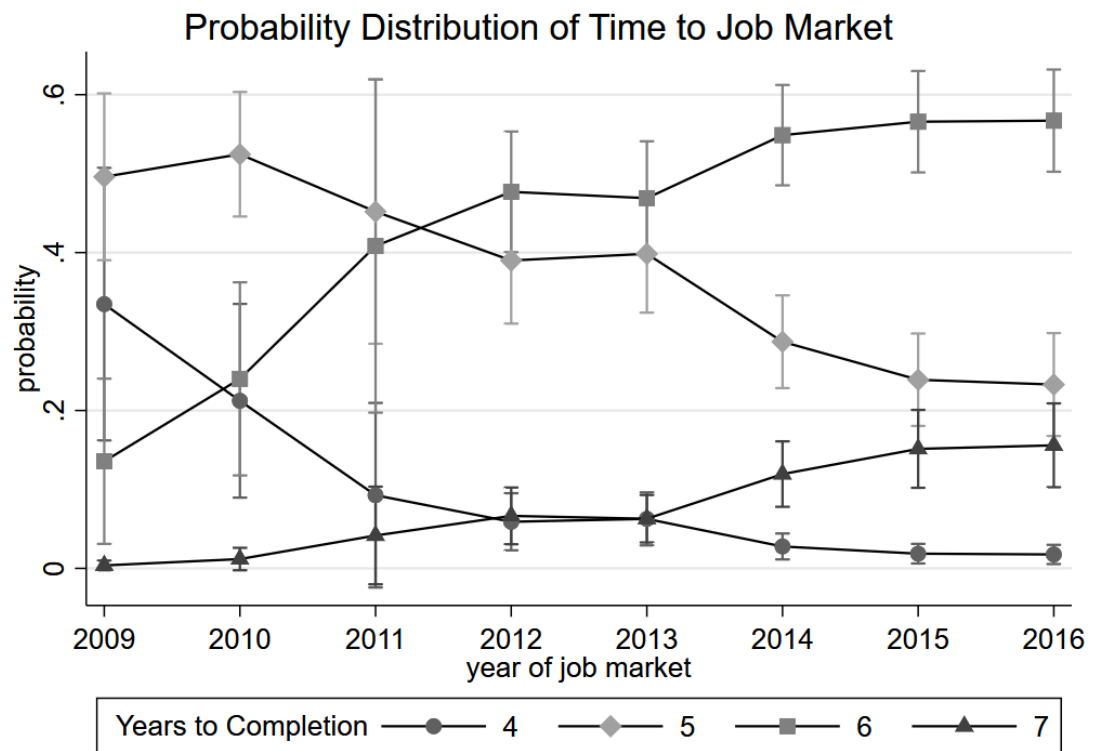


Figure 3.3

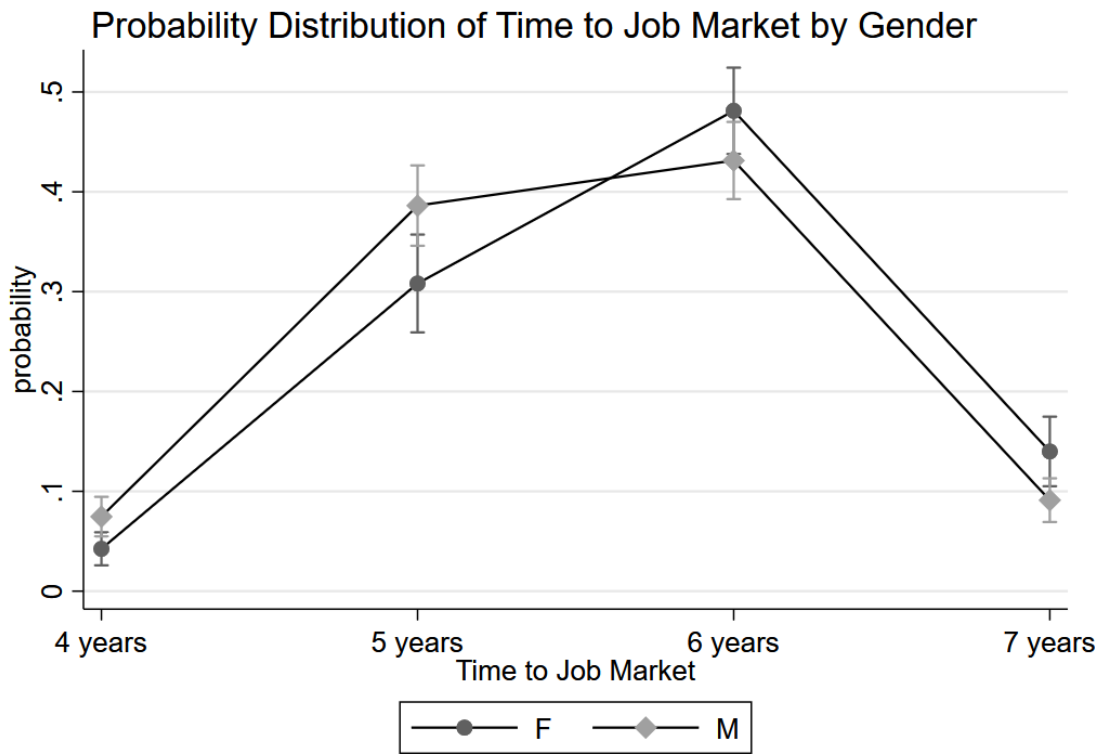
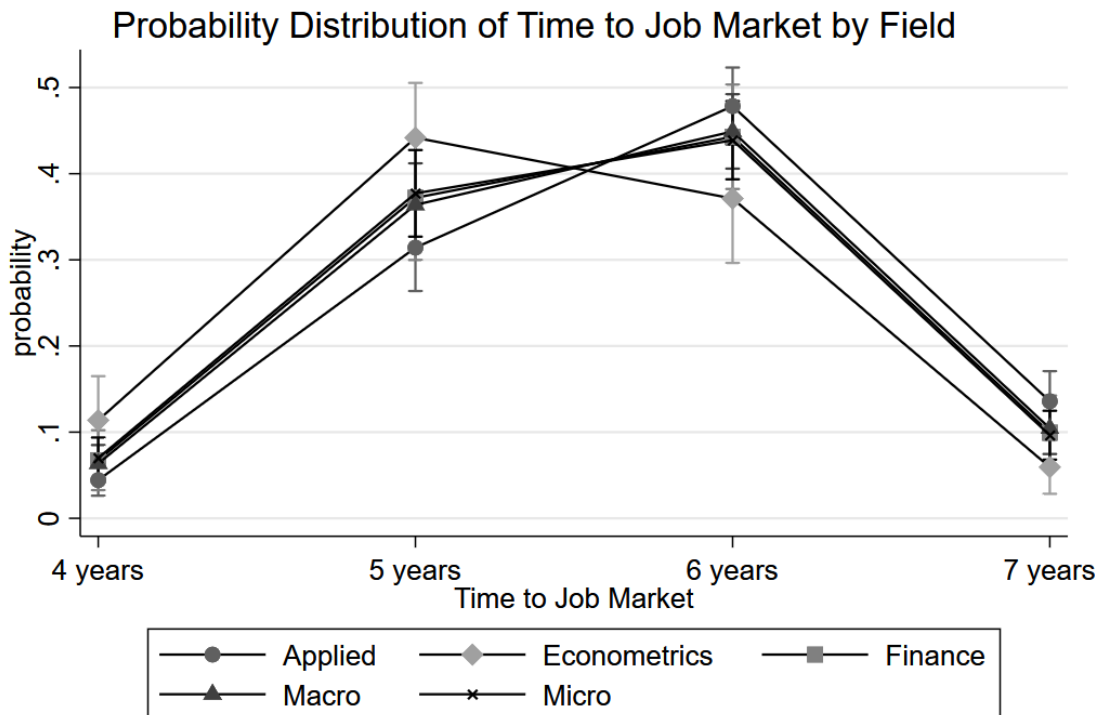
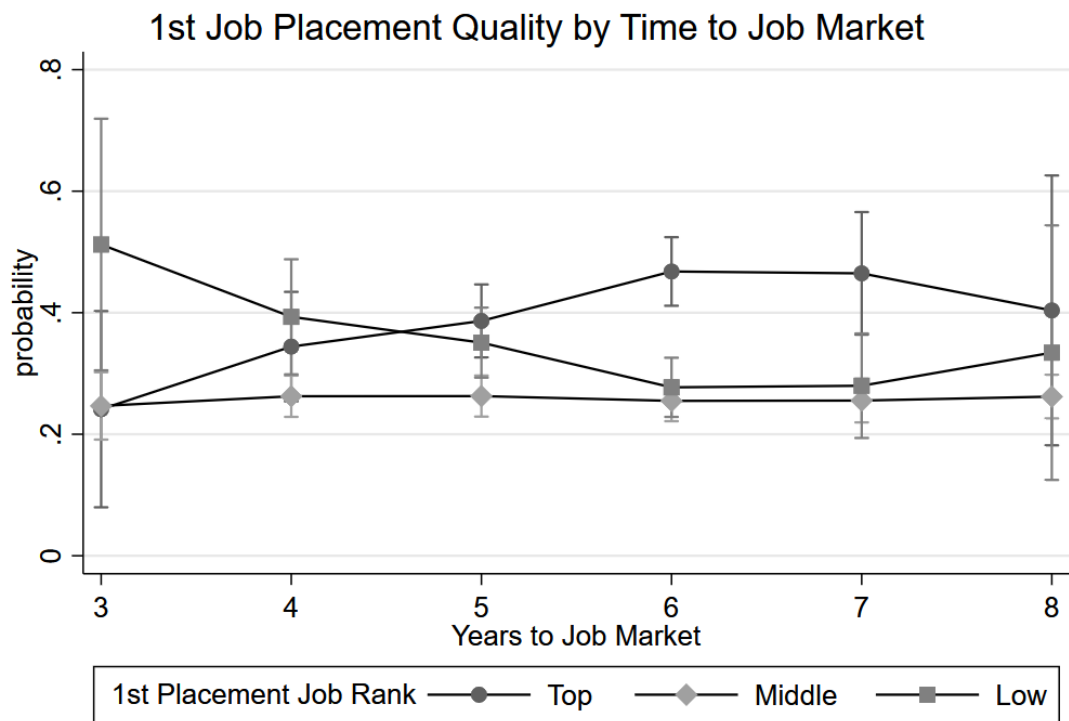


Figure 3.4



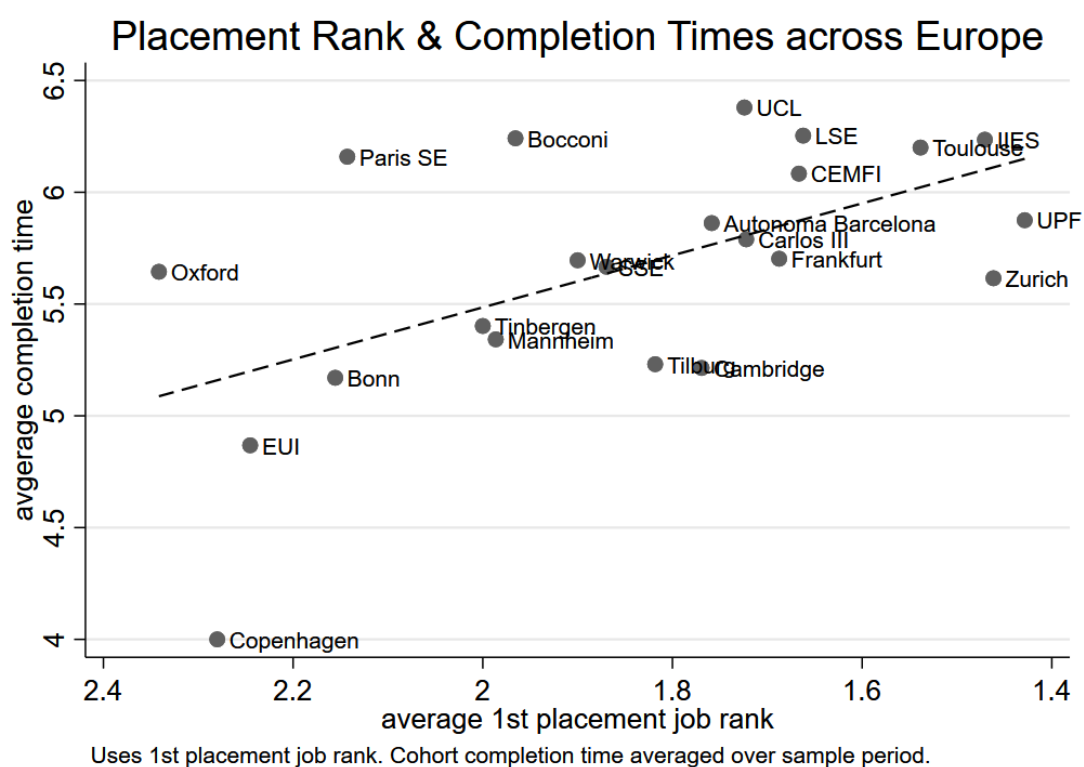
Note: Dots mark point estimates and upper and lower bars mark 95 percent confidence intervals.

Figure 3.5



Note: For details on the Job Ranking, please refer to the text.

Figure 3.6



Note: The figure plots the average completion time against the average rank of the first position for graduates of each PhD-granting institution in the sample.

Figure 3.7

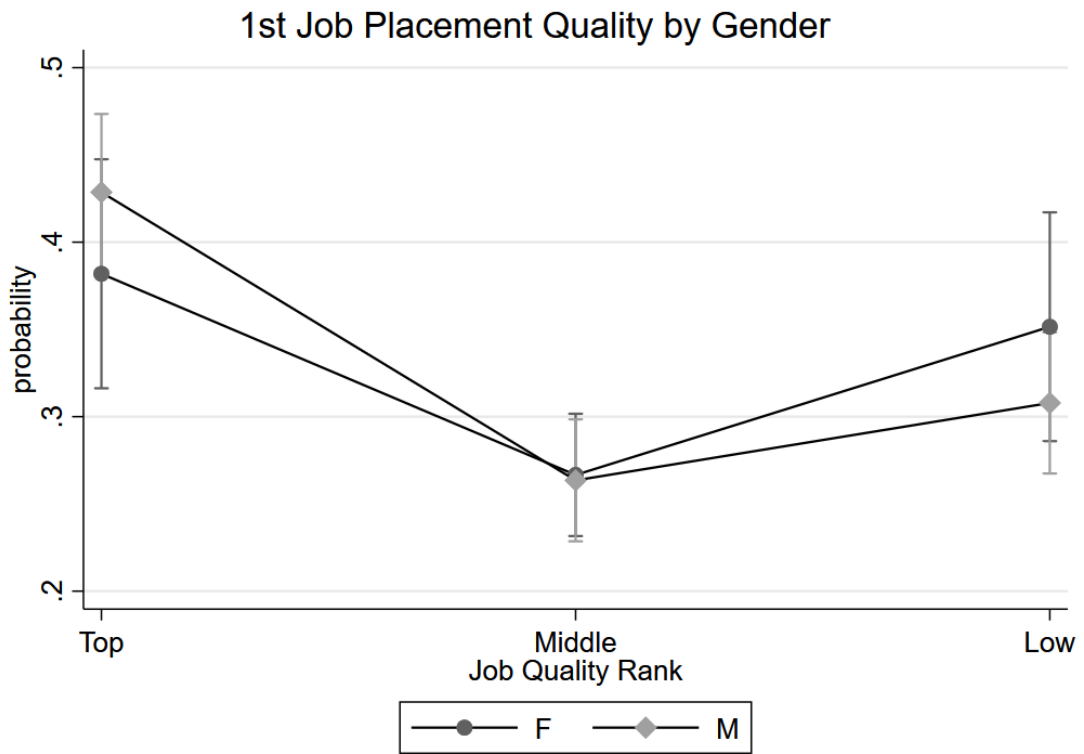


Figure 3.8

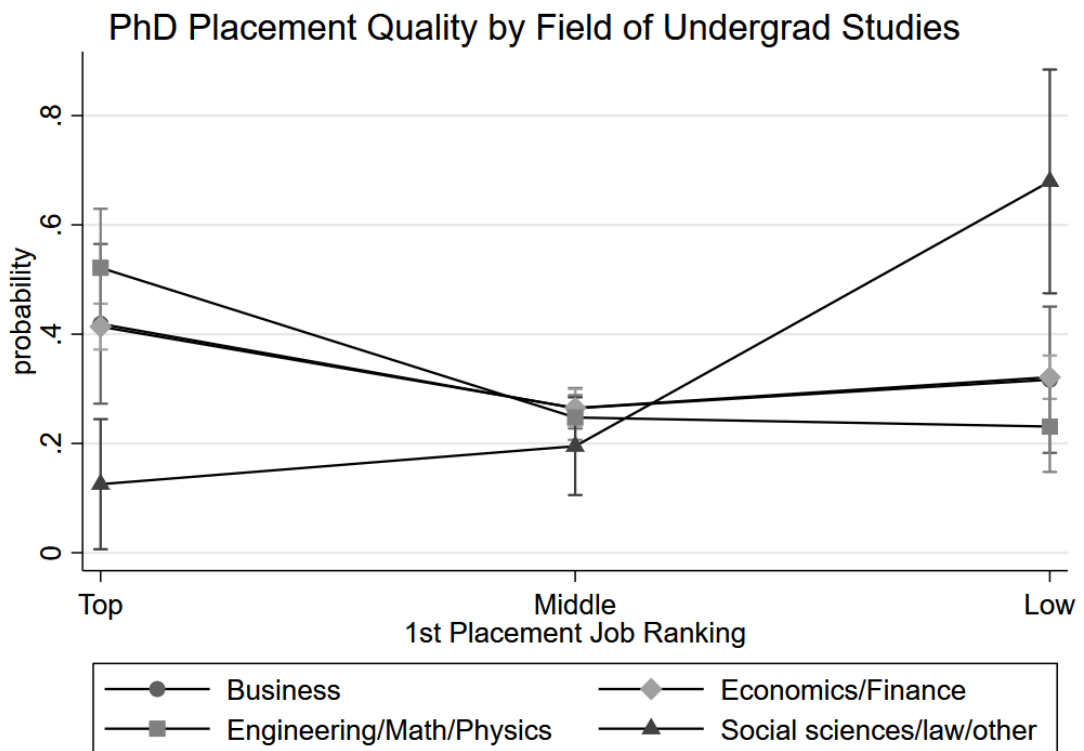


Figure 3.9

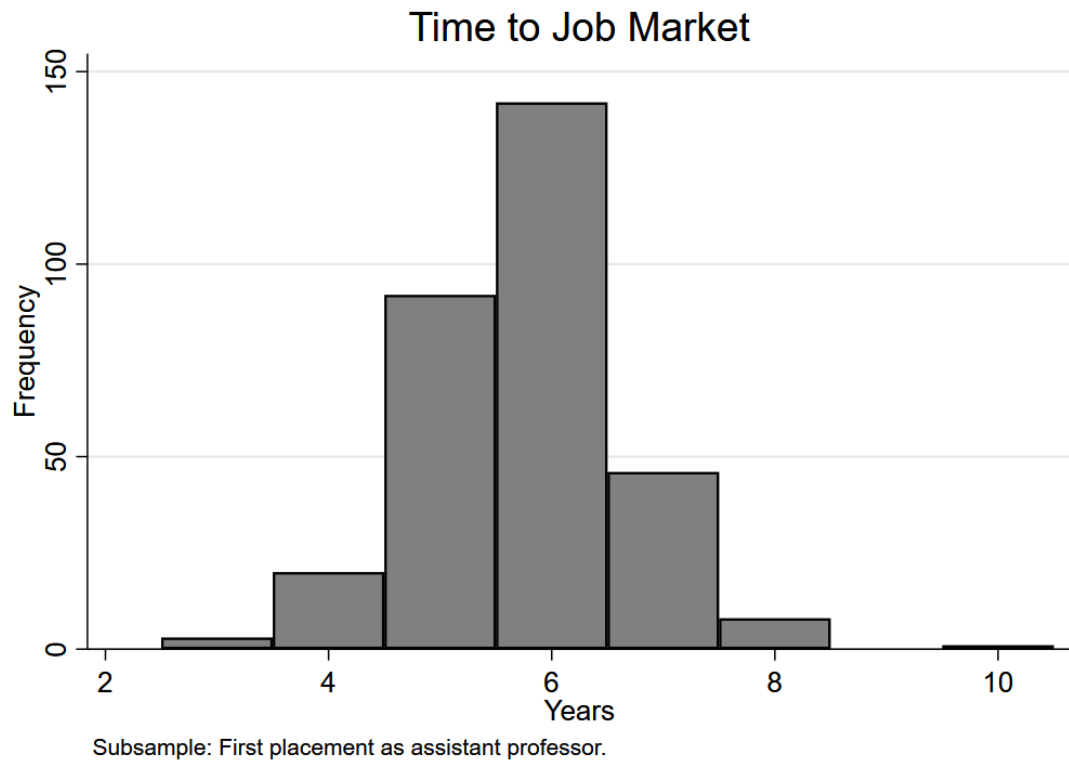
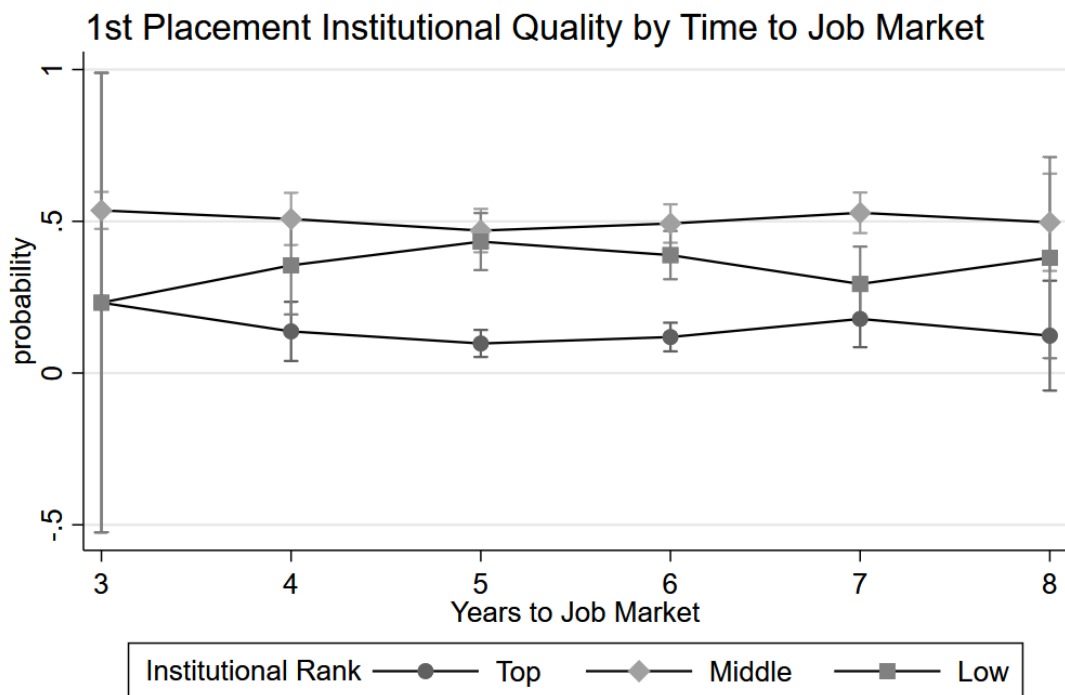


Figure 3.10



For details on Institutional Ranking, please refer to the text.
 Subsample: First placement as assistant professor.

Figure 3.11

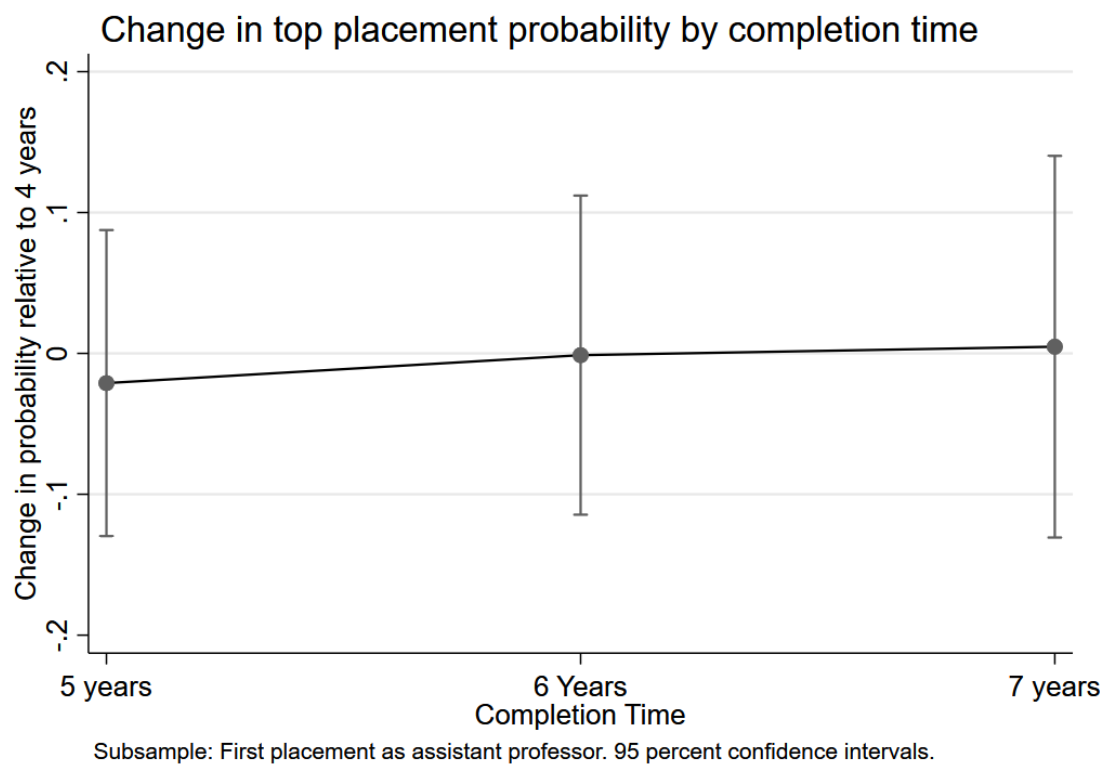
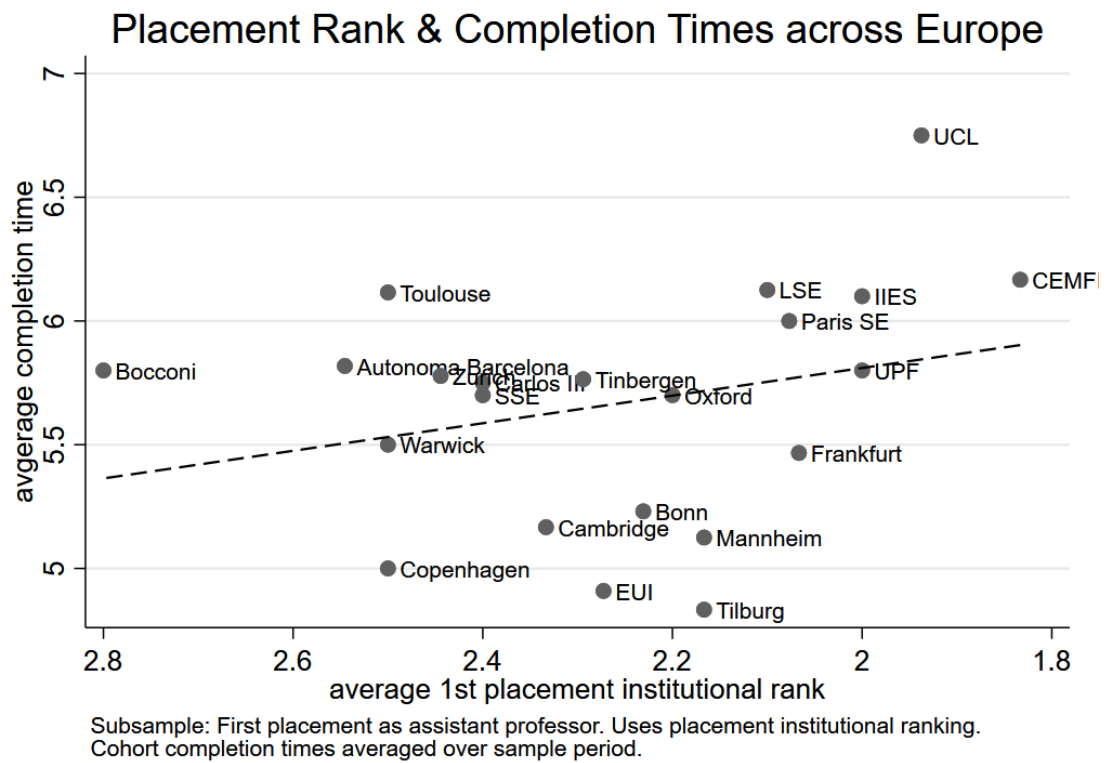


Figure 3.12



Note: The figure plots average completion time against average rank of hiring institution for graduates of each PhD-granting institution in the sample.

Bibliography

- Abbott, B., G. Gallipoli, C. Meghir, and G. L. Violante (2013). Education policy and intergenerational transfers in equilibrium.
- Aguiar, M. and E. Hurst (2007). Measuring trends in leisure: The allocation of time over five decades. *The Quarterly Journal of Economics* 122(3), 969–1006.
- Andrews, I., M. Gentzkow, and J. M. Shapiro (2017). Measuring the sensitivity of parameter estimates to estimation moments. *The Quarterly Journal of Economics* 132(4), 1553–1592.
- Autor, D. et al. (2010). The polarization of job opportunities in the us labor market: Implications for employment and earnings. *Center for American Progress and The Hamilton Project*.
- Autor, D. H. and D. Dorn (2013). The growth of low-skill service jobs and the polarization of the us labor market. *The American Economic Review* 103(5), 1553–1597.
- Ben-Porath, Y. (1967). The production of human capital and the life cycle of earnings. *The Journal of Political Economy*, 352–365.
- Berman, E., J. Bound, and S. Machin (1998). Implications of skill-biased technological change: international evidence. *The quarterly journal of economics* 113(4), 1245–1279.
- Blundell, R., M. Graber, and M. Mogstad (2015). Labor income dynamics and the insurance from taxes, transfers, and the family. *Journal of Public Economics* 127, 58–73.
- Bovenberg, A. L. and B. Jacobs (2005). Redistribution and education subsidies are Siamese twins. *Journal of Public Economics* 89(11), 2005–2035.
- Bowlus, A. J. and C. Robinson (2012). Human capital prices, productivity, and growth. *The American Economic Review* 102(7), 3483–3515.
- Cai, Z. and J. Heathcote (2018). College tuition and income inequality.
- Chetty, R., J. Friedman, E. Saez, N. Turner, and D. Yagan (2017). Mobility report cards: The role of colleges in intergenerational mobility.
- Chetty, R., A. Guren, D. Manoli, and A. Weber (2011). Are micro and macro labor supply elasticities consistent? A review of evidence on the intensive and extensive margins. *The American Economic Review* 101(3), 471–475.
- Chetty, R., N. Hendren, P. Kline, and E. Saez (2014). Where is the land of opportunity? the geography of intergenerational mobility in the United States. *The Quarterly Journal of Economics* 129(4), 1553–1623.
- Ciccone, A. and G. Peri (2005). Long-run substitutability between more and less educated workers: evidence from us states, 1950–1990. *The Review of Economics and Statistics* 87(4), 652–663.
- Coen-Pirani, D. (2017). Geographic mobility and redistribution: A macro-economic analysis.
- College Board (2013). Trends in student aid 2013.
- Consumer Financial Protection Bureau (2012). Private student loans. Technical report.
- Cortes, G. M., N. Jaimovich, and H. E. Siu (2017). Disappearing routine jobs: Who, how, and why? *Journal of Monetary Economics* 91, 69–87.

- Dwyer, R. E. (2013). The care economy? gender, economic restructuring, and job polarization in the us labor market. *American Sociological Review* 78(3), 390–416.
- Epple, D., R. Romano, S. Sarpça, and H. Sieg (2013). The US market for higher education: A general equilibrium analysis of state and private colleges and public funding policies.
- Epple, D., R. Romano, and H. Sieg (2006). Admission, tuition, and financial aid policies in the market for higher education. *Econometrica* 74(4), 885–928.
- FinAid (2016, November). Historical loan limits.
- FinAid (2018, February). Historical interest rates.
- Findeisen, S. and D. Sachs (2016a). Education and optimal dynamic taxation: The role of income-contingent student loans. *Journal of Public Economics* 138, 1–21.
- Findeisen, S. and D. Sachs (2016b). Optimal need-based financial aid.
- Fuller, M. B. (2014). A history of financial aid to students. *Journal of Student Financial Aid* 44(1), 42–68.
- Gale, W. G. and J. K. Scholz (1994). Intergenerational Transfers and the Accumulation of Wealth. *The Journal of Economic Perspectives* 8(4), pp. 145–160.
- Goos, M., A. Manning, and A. Salomons (2014). Explaining job polarization: Routine-biased technological change and offshoring. *The American Economic Review* 104(8), 2509–2526.
- Güvenen, F. and B. Kuruscu (2010). A quantitative analysis of the evolution of the us wage distribution, 1970–2000. *NBER Macroeconomics Annual* 24(1), 227–276.
- Güvenen, F. and B. Kuruscu (2012). Understanding the evolution of the us wage distribution: A theoretical analysis. *Journal of the European Economic Association* 10(3), 482–517.
- Güvenen, F., B. Kuruscu, and S. Ozkan (2014). Taxation of human capital and wage inequality: A cross-country analysis. *The Review of Economic Studies* 81(2), 818–850.
- Haider, S. J. (2001). Earnings instability and earnings inequality of males in the united states: 1967–1991. *Journal of Labor Economics* 19(4), 799–836.
- Hanushek, E. A., C. K. Y. Leung, and K. Yilmaz (2003). Redistribution through education and other transfer mechanisms. *Journal of Monetary Economics* 50(8), 1719–1750.
- Havranek, T. (2013). Publication Bias in Measuring Intertemporal Substitution. *Working Papers IES* 2013(15).
- Havránek, T. (2015). Measuring intertemporal substitution: The importance of method choices and selective reporting. *Journal of the European Economic Association* 13(6), 1180–1204.
- Heathcote, J., K. Storesletten, and G. L. Violante (2017). Optimal tax progressivity: An analytical framework*. *The Quarterly Journal of Economics* 132(4), 1693–1754.
- Hendricks, L. and O. Leukhina (2017). How risky is college investment? *Review of Economic Dynamics* 26, 140–163.
- Herrington, C. M. (2015). Public education financing, earnings inequality, and intergenerational mobility. *Review of Economic Dynamics* 18(4), 822–842.
- Holmlund, H., M. Lindahl, and E. Plug (2011). The causal effect of parents’ schooling on children’s schooling: A comparison of estimation methods. *Journal of Economic Literature* 49(3), 615–51.
- Holter, H. A. (2015). Accounting for cross-country differences in intergenerational earnings persistence: The impact of taxation and public education expenditure. *Quantitative Economics* 6(2), 385–428.
- Hoxby, C. M. (2015). Computing the value-added of american postsecondary institutions.
- Hoxby, C. M. (2016a). The dramatic economics of the u.s. market for higher education. *NBER Reporter* 3.
- Hoxby, C. M. (2016b). The productivity of u.s. postsecondary institutions: Implications for markets and policies. The Fellows Lecture, The 2016 Annual Meeting of the Society of Labor Economists.
- Hsieh, C.-T., E. Hurst, C. I. Jones, and P. J. Klenow (2016). The allocation of talent and us economic

growth.

- Huggett, M., G. Ventura, and A. Yaron (2011). Sources of lifetime inequality. *The American Economic Review* 101(7), 2923–2954.
- Institute for Higher Education Policy (2003). Private loans and choice in financing higher education.
- Jacobs, B. and A. L. Bovenberg (2011). Optimal taxation of human capital and the earnings function. *Journal of Public Economic Theory* 13(6), 957–971.
- Johnson, N. (2014). College costs and prices: Some key facts for policymakers. *Lumina Issue Papers*.
- Katz, L. F. and K. M. Murphy (1992). Changes in relative wages, 1963–1987: Supply and demand factors. *The Quarterly Journal of Economics*, 35–78.
- Keane, M. and R. Rogerson (2012). Micro and macro labor supply elasticities: A reassessment of conventional wisdom. *Journal of Economic Literature* 50(2), 464–76.
- Kotera, T. and A. Seshadri (2017). Educational policy and intergenerational mobility. *Review of Economic Dynamics* 25, 187 – 207.
- Krueger, D. and A. Ludwig (2016). On the optimal provision of social insurance: Progressive taxation versus education subsidies in general equilibrium. *Journal of Monetary Economics* 77, 72–98.
- Landersø, R. and J. J. Heckman (2017). The Scandinavian fantasy: The sources of intergenerational mobility in Denmark and the US. *The Scandinavian Journal of Economics* 119(1), 178–230.
- Lee, S. Y. T. and A. Seshadri (2014). On the intergenerational transmission of economic status.
- Lindenlaub, I. (2017). Sorting multidimensional types: Theory and application. *The Review of Economic Studies* 84(2), 718–789.
- Lochner, L. and A. Monge-Naranjo (2011a). Credit constraints in education.
- Lochner, L. J. and A. Monge-Naranjo (2011b). The Nature of Credit Constraints and Human Capital. *The American Economic Review* 101(6), 2487–2529.
- McDaniel, C. (2007). Average tax rates on consumption, investment, labor and capital in the OECD 1950–2003.
- MeasureOne (2013). Private student loan report 2013.
- Meghir, C. and L. Pistaferri (2004). Income variance dynamics and heterogeneity. *Econometrica* 72(1), 1–32.
- Mulligan, C. B. (1999). Galton versus the human capital approach to inheritance. *Journal of Political Economy* 107(S6), S184–S224.
- National Center for Education Statistics (2017). The condition of education.
- Sachs, D., A. Tsyvinski, and N. Werquin (2016). Nonlinear tax incidence and optimal taxation in general equilibrium.
- Stantcheva, S. (2017). Optimal taxation and human capital policies over the life cycle. *Journal of Political Economy* 125(6), 1931–199.
- Stinebrickner, R. and T. Stinebrickner (2008). The effect of credit constraints on the college drop-out decision: A direct approach using a new panel study. *The American Economic Review* 98(5), 2163–2184.
- Stock, W. A., T. A. Finegan, and J. J. Siegfried (2009). Completing an economics phd in five years. *The American Economic Review* 99(2), 624–629.
- Stock, W. A. and J. J. Siegfried (2014). Fifteen years of research on graduate education in economics: What have we learned? *The Journal of Economic Education* 45(4), 287–303.
- Storesletten, K., C. I. Telmer, and A. Yaron (2004). Cyclical dynamics in idiosyncratic labor market risk. *Journal of Political Economy* 112(3), 695–717.
- Teulings, C. N. (2005). Comparative advantage, relative wages, and the accumulation of human capital. *Journal of Political Economy* 113(2), 425–461.

Yamaguchi, S. (2012). Tasks and heterogeneous human capital. *Journal of Labor Economics* 30(1), 1–53.